THE RE-ENGINEERING OF VHF MOBILE RADIO SERVICES IN THE UK

Roger Edward Fudge

Submitted for the degree

Doctor of Philosophy

THE CITY UNIVERSITY

Department of Electrical and Electronic Engineering

April 1984

Volume I

TABLE OF CONTENTS

TABLE OF CONTENTS

TABLE OF CONTENTS

TABLE OF CONTENTS

Page Number

TABLE OF CONTENTS

TABLE OF CONTENTS

TABLE OF CONTENTS

TABLE OF CONTENTS

TABLE OF CONTENTS

Page Number

TABLE OF CONTENTS

TABLE OF CONTENTS

TABLE OF CONTENTS

List of Figures

PAGE

NUMBERING

AS ORIGINAL

List of Figures

List of Figures

List of Figures

List of Figures

List of Figures

List of Figures

List of Figures

List of Figures

List of Figures

ACKNOWLEDGEMENTS

The acknowledgements I should like to make for support for this task are many and varied. They are given below in no particular order of merit.

To the Home Office (Directorate of Telecommunications) for permission to undertake the work, providing the financial resources, and granting me the facilities to comply with at least the minimum attendance requirement of the University. To the City University, and in particular Professor A C Davies, for accommodating this work and for his overseeing support and encouragement. To Mr W S Wilkinson, who as my internal supervisor has spent many hours in valued discussion and guidance. To Mr R M Hughes, who as my external supervisor in the Directorate also provided encouragement, understanding, and was in a position for the early stages, to ensure that this project took its place with other official demands.

This project did not start from 'cold' but benefited from experience gained with forerunning work. In that, and during the project, I was fortunate to have a team of support staff under my direction. This team was at best small, and was subject to continual change. I should like to thank the following for their contributions to the tasks which were largely on the experimental and detailed analysis aspects. They are in general credited where possible in the text with their contributions, where this is possible, via references. But here I wish to give their names with the topic with which they were associated. A N Kent (early work on modulation subjective trials), V Asque (later work on modulation subjective trials), P J Bridgeman (modulation comparison trials and initial intermodulation work), B D Hill (modulation comparison trails and initial intermodulation work), R S Keeble (intermediate intermodulation work), L D Clark (later intermodulation work), J R Pearce (frequency planning and Rayleigh fading work [see references]).

In the final, but of course not least, category comes my family who have had to 'make do' without me for at least the last year of detailed writing. My father, E Fudge, for assistance in typing correction, my sons G R and D E Fudge for typing corrections and standing-in for my absence in general. To my wife Eileen who has tolerated being a 'thesis widow' and excused me from such time consuming tasks as washing-up, gardening, decorating etc.

DECLARATION

ABSTRACT

This thesis covers the work and considerations necessary for the design of VHF mobile radio systems, and is directed particularly at the requirements of the emergency services of the United Kingdom. It was occasioned by an international regulatory decision to re-allocate part of their existing operating frequencies. This work provides both a structure for such a design, and a record of events. It considers the various elements of design, initially from a theoretical base which is developed to the extent where practical trials and external constraints overtake the theoretical analysis. The constraints can take the form of compatibility with the existing system, cost, timescale, and political considerations. The concept of a design flow chart is shown to lead to many circular arguments of the chicken-and-egg variety to the extent that the resultant diagram is termed in interaction diagram since its main function is to show how the design elements interface with each other. It consists of many feed-back loops; and design completion is achieved when a traverse of all paths of the interaction diagram can be made without meeting discontinuities at the interfaces.

The topic of area coverage is examined in depth and the multiple transmitter common channel (quasi synchronous) mode is shown to be efficient from both the time and the spectrum economy viewpoints. There are benefits also from operational considerations. The section shows that a reuse factor of 3 should be practical, which is of interest to mobile radio users in general, and cellular radio telephony systems in particular. The choice of type of modulation is quickly shown to lie between FM, AM and SSB. All three are analysed from single station and quasi synchronous aspects: the theoretical predictions of the superiority of AM, for the system under study, is confirmed by subjective trials from which design parameter limits are obtained. Examination of equipment hardware, mobile and base station, shows that system design aspects are severely constrained by intermodulation products due to passive site non linearities. Measurements indicate that these will be troublesome up to at least the 11th order. A frequency assignment plan covering mobile and fixed services for the country is derived and the dominating influence of the intermods is shown to result in a workable, but unconventional, version.

GLOSSARY OF TERMS AND ABBREVIATIONS

ALC:	Automatic level Control.
AM:	Amplitude modulation.
AVL:	Automatic vehicle location.
Base Station:	The location of the fixed radio equipment (tranmitters, receivers, etc) which interfaces with the mobile radio equipment. (Synonymous with Hill Top Site).
Beat Cycle:	One cycle of the frequency difference between two carriers.
Control (Room):	The fixed terminus of the mobile radio communications chain.
Common Channel:	A multiple tranmitter area coverage technique where all the tranmitter frequencies lie within one RF channel.
DE:	De-emphasis.
Directorate:	Directorate of Telecommunications of the Home Office
Fixed Link:	See Link.
Fixed Segment:	The equipment and channel(s) which interface with the control (including repeated links).
FM:	Frequency modulation.
Forward:	In the direction of control to mobile = outward/outgoing.
Hill Top Site:	See base station.
HPA:	High power amplifier.
IF:	Intermediate frequency.
Incoming:	See Return.
Intermod:	Intermodulation product.
IPA:	Intermediate power amplifier.
Link:	The fixed highway between a control and base station between two base stations.
LPA:	Low power amplifier.
LPC:	Linear predictive coder.
Master Station:	A base station which carries ordinary and repeated links.

Mobile:	The mobile terminus of the mobile radio communications chain - a vehicle or its radio equipment.
Mobile Segment:	The equipment and channels which interface with the mobile.
Modem :	Modulator/demodulator, generally for digital working on audio channels.
Mute:	A devise for inhibiting a receiver's audio output when this would be unacceptable = squelch.
Mute Tail:	The burst of noise at the output of a receiver when the input has dropped to a low level but the mute time constants have not disabled the audio output.
Outward:	See Forward.
PCM:	Pulse code modulation.
PE:	Pre-emphasis.
PEP:	Peak envelope power.
Pip-Tone:	An 'engaged' signal comprising a discontinuous tone modulation transmitted by base station(s) when a mobile is using the return part of the channel.
PM:	Phase modulation.
Quasi Synchronous:	A technique of area coverage using multiple transmitters operating on very nearly the same radio frequency.
QS:	Quasi synchronous .
Repeated Link:	A fixed link from one base station to another.
Repeater Station:	A base station which carries ordinary and repeated links = master station.
Return:	In the direction from mobile to control = incoming/inward.
RF:	Radio Frequency.
SINAD:	Signal In Noise And Distortion - a measure of quality.
Spaced Carrier:	A technique of area coverage using multiple transmitters operating on significantly different frequencies but all within one RF channel bandwidth.

Sporadic E:	A form of anomolous propagation caused by reflections in the E layer of the Ionosphere.
Squelch:	See Mute.
SS:	Spread spectrum (modulation).
SSB:	Single side band (modulation).
VLSI:	Very large scale integration.
WARC(79):	World Administrative Radio Conference - 1979.

SYMBOLS

Note:- i) Some symbols have suffixes to denote various subsets but these are not used where there is no ambiguity.

ii) Some symbols are used for more than one meaning, the various versions are shown, and the one to use will be obvious from the local definition and context.

a	A distance variable
A	Address time
An	Area
b	Distance variable
B	RF Bandwidth
c	Distance variable
C	Carrier power
d	Distance variable, specifically from edge of coverage or between cells
Em	Field strength at a mobile
Er	Voltage at the input to a mobile receiver
f	Frequency, specifically the highest baseband frequency
Fn	Frequency
F(d)	A propagation law
ΔF	Peak frequency deviation
G	Gain
Gn	Gain of an aerial
$G(\theta,\alpha)$	Gain of an aerial as a function of angle θ or α
h_n	Height
H	Threshold voltage, or form of coupling harness
i	Distance variable, specifically a coordinate in hexagonal geometry
I	Integer variable
j	Distance variable, specifically a coordinate in hexagonal geometry
J	Proportionality constant, or an integer
K_n	Constants
K_p	Propagation constant for p % of locations.
L	Length
m	Deviation ratio $\frac{\Delta F}{f}$, modulation index, number of elemental cells

M	Message time
n	A general number of quantity
n	Baseband noise power
N	RF noise power
No	Noise power density
P_x	Power
P	Carrier amplitude from base station P
Pn	Power of the noise
q	Distance variable
Q	Carrier amplitude from base station Q, query time
r	Distance variable, range bandwidth
R	Resultant carrier amplitude, response time
s	Distance variable, bandwidth
s/n	Baseband (audio) signal to noise ratio
Sr	RMS value of Rayleigh distribution
Ss	RMS value of Suzuki distribution
S/N	RF signal to noise ratio
t	Time, a distance variable, bandwidth
Tn	Time
u	Reuse distance
U	20 Log_{10} e
V	Instantaneous emf, vehicle velocity
Vc	Voltage of carrier
Vn	Voltage of noise
w	Bandwidth
x	Distance variable, bandwidth
y	Distance variable, bandwidth
z	Parameter of Susuki distribution
Z	Switching range (number of channels to be available)

α Angle variable

β Angle variable

ΔF Peak frequency deviation

Γ Protection ratio

Γ' Modified protection ration = $\Gamma\left(\frac{\Lambda}{1-\Lambda}\right)$

Γ_a Protection ratio for AM

Γ_f Protection ratio for FM

$\Gamma(\ )$ Gamma function

θ Angle variable

λ Wavelength

σ Distribution parameter, specifically of Raleigh distribution function

τ_n Fractional time or probability

ϕ Angle variable

ω Angular frequency, specifically as a difference frequency

Ω Angular frequency, specifically as a resultant

Λ Probability that the field strength due to one tranmission is greater than another by Γ

1. Introduction

This thesis is concerned with the design of a Land Mobile Radio System.

Specifically it relates to the design and realisation of a mobile radio system for the emergency services of England and Wales. The opportunity to actively consider the design of such a complete and complex system in its entirety will be rare* and it is necessary to explain how this arose, my position in the responsible organisation and the type of thesis which results.

The Directorate of Telecommunications of the Home Office is responsible for the communications aspects of the Police and Fire Emergency services for England and Wales, its statutory responsibilities are in the advisory, regulatory, and fiscal areas. This has developed however to cover also the implementation of the bulk of the radio schemes for these users - comprising the purchase, installation, maintenance and recovery of rental charges for a very large quantity of equipment. The systems in use have been developing slowly in response to demands from the users and advances in techniques; the background section describes the types of system in use currently and some of the evolution - all of which is pertinent to this study. A discontinuity was imposed on this evolution by the decision of the World Administrative Radio Council in plenary session in late 1979 (WARC 79). This body, operating under United Nations auspices, reassigned certain radio frequencies the consequences of which were that half the (already crowded) frequencies allotted to the emergency services were to be given over to Broadcasting Services in both the UK and on the Continent by the end of 1989. Thus new bands of frequencies needed to be located and the design of the system for these new bands could be considered from scratch.

Since the responsibility for writing the specifications for new equipment lay within the Research and Development Division, I was

---

* A possible exception is the new mobile radio telephony system (the so called cellular system) currently under consideration but for that many design aspects have already been considered in the USA and elsewhere

involved in the system design as head of the Radio Systems Strategies Section in order to define the system against which the individual equipments could be specified and in a consultative role for the details of these specifications. A reorganisation of the Directorate someway through this period of study abolished a specific R&D Division (which covered all aspects of telecommunications) and left me as head of Mobile Radio Research in the Mobile Radio Division, my role continued as before however. The small supporting team under my control became smaller at the same time. Virtually all theoretical studies and system concepts were undertaken by me with experimental and trials work conducted by the team under my control and additional inputs coming from contracts for specific tasks placed with Industry or Universities and under my control. The non personal contributions to this work will be flagged at the appropriate points.

This thesis will have a different form to those normally submitted for this degree. It will not cover one particular topic in great depth but rather will seek to develop the theoretical basis necessary for an understanding of both the elements of this relatively complex system and their interactions. But an important part of the thesis will be to show the limit to which it is sensible to take the theoretical considerations before undertaking experimental work and the scientific factors start to be outweighed by considerations of implementation in a real world situation. These include the constraints of:- finances, manpower resources, time-scale, and political considerations. It is hoped that the thesis will provide both a record of the work that was performed in this respect and also a kind of template against which either similar systems could be designed or more importantly against which changes to this design can be evaluated.

2. Background

2.1 The Starting Point

The design of a mobile radio system starts with the requirements of the user, in this case a police force or fire brigade. Put very briefly this requirement is to provide a radio scheme which enables two way communications to be achieved from one or more fixed control locations to one or more of a fleet of mobile vehicles (mobiles), and to permit mobile to mobile (car to car) communications in a manner under the control of a controller. The area to be covered is that over which the controller exercises an influence, this is usaully a geographic county, and there will be requirements to cover parts of a county or even very localised regions such as may come under the influence of say one police station in a town. This latter requirement is usually met by a UHF low powered system which caters for both personal (hand-held) radios and vehicle mounted versions; this type of operation is not considered here and, due to the way in which it developed, it has very little impact on the mobile system which has traditionally been at VHF.

A simple and idealised diagram of a typical VHF mobile communications system, of the type under discussion, is shown in figure 2.1.1. The headquarters control can reach a mobile by a two-hop path. That from control to a hill top base station, is via a fixed radio link usually operating at VHF. The signal is demodulated at the base station and remodulated onto another VHF carrier to access the mobile. Depending on the mode of operation, one or more base stations may convey the message. The return path from the mobile is the inverse of the outgoing path but uses different frequencies and different equipment. This very simple description is expanded in the next and subsequent chapters.

The degree of cover in the area of operation is perhaps best described as 'that which is most cost effective'. This rather begs the issue and makes rigorous design difficult but it does allow flexibility in practice so that the users who can afford more money can expect a better system whilst those who may feel their system inadequate can express their dissatisfaction and willingness to overcome it in financial terms. A quantified version of the requirement has been postulated as intelligible communications for at least 90% of locations in the coverage area for at least 99% of the time. For the location part of this, interest is concentrated on the areas where communications are most frequently needed. ie. centres of population, transport routes, holiday spots etc.

The rather loose specification was largely a result of the way in which the mobile communication system had developed. Originally two frequency bands were allotted for the emergency services mobile operations (the mobile segment), one for the return path and one for the outgoing, and these have, after some relatively minor alterations, become approximately 80-84 MHz and 100-104 MHz respectively. Originally channel spacings of 200 kHz were used and the area of operation was defined by the transmitter powers, aerial heights and locations, and receiver performances then in use. These factors were modified over the course of time until a user was satisfied with the situation. The channel spacing was halved from time to time and the degree to which a channel was re-used by another user elsewhere increased in order to cope with ever increasing demand for more channels. Thus the system was able to grow and evolve slowly with only small changes made at a time and sufficient time to adapt these into workable ones.

Amplitude modulation was used and where there was a demand for area coverage (see the section on area coverage) this was met by a technique called spaced carrier operation. The last reduction in channel bandwidth from 25 kHz to 12.5 kHz was made in 1978 and a few systems operating on a quasi synchronous method of area coverage were introduced. Since the control was unlikely to be favourably located so far as radio propagation was concerned and since in any case a number of transmitting/receiving sites (base stations) were required to cover a county, the control was linked to the base stations by means of fixed VHF radio links (often termed the link) operating in the bands 146-148 MHz and 154-156 MHz.

2.2 Summary of Initial Factors

The situation of the Directorate at the time of WARC 79 can be summed up, so far as the mobile radio communications operations were concerned, as:-

- Regulatory control of the 43 counties of England and Wales

- An average of some 5 channels needed per county

- 4 + 4 MHz available for mobile communications plus
  2 + 2 MHz for fixed link feeder services (and special facilities)

- Effectively direct control of police communications covering 40 counties and the bulk of the fire brigades.

- Effective control for these bodies of some 200 base station hill-top sites.

- Teams of engineers and technicians skilled at installing and maintaining a slowly evolving system.

- Design engineers having to update and enhance the system whilst maintaining continuity of service, which tended to piecemeal operations conducted on a trial basis.

- Severe congestion in the use of the link frequencies (in effect 200 two-way channels carried in 1 + 1 MHz) and congestion in the, mainly 25 kHz channelled, mobile bands.

The decision of WARC 79 to assign the band 100-104 MHz to broadcasting inside 10 years thus created at once both a serious problem for re-design and re-equipping, and a rare opportunity to virtually start afresh with a planned integrated system which would allow for future growth and the incorporation of future developments in a tractable manner.

2.3 Time Factors

Experience with a continual round of purchasing and re-equipping which is necessary to cope with time expired equipment and growth, established that 5 years was the minimum which should be allotted to the physical changeover for the equipment (this allowed a 1 year contingency). The projected start of physical changeover was then the start of 1985. A one year allowance from the date of order was made for the manufacture of the equipment for this first phase (this is a representative figure - experience has shown that some elements may well require longer). The Directorate has a policy of evaluating equipment before purchase and the evaluation of a range of equipments was projected to take 1½ years; an allowance of 6 months was deemed necessary for contractual matters associated with both the main purchase and the evaluation models; the time to manufacture the evaluation models was set as one year. Thus detailed equipment specifications were required by mid 1980.

Prior to this it was necessary to consider and obtain (a) new frequency band(s), consider the design of the system which had to rely on devices and techniques proven by the date of the first evaluation equipment but which would be being installed as new equipment some 8-9 years later, optimise the system design, conduct research/evaluation on any uncertain aspects, and write the detailed specification for some 40 items of equipment. Clearly an impossible task. A bar chart of the time scales is shown in figure 2.3.1.

The timescale therefore effectively decreed that the services to be provided to the user would only be those already offered; this excluded a number of possible enhancements, notable among which was medium to high speed data transmission (anything above some 300 bit/sec already in use for rudimentary resource availability systems and implemented as tones on the speech channel from mobiles only). Thus data or speech-plus-data systems were not to be considered.

My initial tasks were to provide advice on the choice of frequency bands, advise on the choice of modulation, and advise on equipment specification details. The modulation to be used was considered an open question since either improved performances might be obtained (compared to the AM system in use), time scales might be shortened, and/or costs reduced by adopting a standard already in use elsewhere. There were additional factors of a more political nature which will be covered in the relevant section. Any research associated with these tasks could take no more than six months.

It was obvious that the operating parameters and hence the specification for any one piece of equipment depended on the parameters of both the systems in total and the parameters of the pieces of equipment adjacent to it in the communications chain. For this reason, if no other, it would be necessary to consider equipment design as a part of the system design with a two-way inter-action between them. Therefore the study to be described here was undertaken as a means of attempting to satisfy both the stated short term requirements and as a means of providing support to a project which although desirably serial in nature would have to operate as a parallel exercise with incomplete intermediate cross links between the parallel paths. To be able to operate such an overall project as that demanded by WARC 79 on this basis would require the fullest possible understanding of the technical issues involved.

This thesis therefore attempts to cover those issues for the design of a mobile radio system for general use but with particular reference to those aspects which affect operations for the emergency services and against the background stated here.

3. Framework for the Systems Design

3.1 Basic System Definition

It is necessary to define the system for which the design is to be done. It will be shown later in this section that there is no real starting point for this exercise, the arguments and concepts are circular ones. Since, however, the need here is fundamentally to change an existing system and that system has the same core form as a number of other mobile radio schemes, this core will be used to define the basic systems.

Figure 2.1.1 is a pictorial representation of such a basic system for the requirements described in the Background section (chapter 2). A county is served from a number of base stations each of which radiates one, or probably more, channels to fire and/or police mobiles. Each base station is connected to control rooms by radio links. Some outgoing channels may be effectively unique to that base station and serve only the area within its range, other channels may be needed to cover greater areas up to that of the 'county' (the term county here is rather a loose one in that some users cover more than one county eg. Devon and Cornwall Constabulary). Base stations may be shared by users in more than one county if their coverage is beneficial, thus base stations will in general be required to handle more channels than that assigned to a county. The existing system has one fixed link carrier from control to each base station for each mobile channel (ie. multiplexed links are not used) and with the desire to minimise changes this was taken as a starting point for the study. Any radio frequency assignments (frequencies) made for one base station, whether mobile or fixed, should be available for re-use elsewhere in England and Wales with the minimum separation consistent with overall optimisation.

A typical county has an area of some 2500km^2 (say 50 x 50km or 30 miles by 30 miles) with four base stations. Since the base stations will seldom be on the edge of the county then approximately omnidirectional coverage will be required.

Simplex operation is required - but this needs qualifying. The basis of operation is of the form:- press to talk, talk, end of message/over. Efficient spectral use has led to two frequency operation with all like users being clustered together - thus all base station transmitters occupy the 100-104 MHz band and all mobile transmitters (base station receivers) occupy the 80-84 MHz band. Since a mobile cannot directly hear whether another is using their shared channel, the base station transmits an 'engaged signal' (a pip-tone) as soon as it starts to receive transmission from a mobile, all mobiles on that channel (except the one transmitting) receive the pip tone and know that access to the channel is barred to them - thus conflict is avoided. This means that whilst the mobiles can operate in a simplex mode, which is cost beneficial, the base station must operate duplex. This often called half duplex.

The foregoing is thought to be a statement of the starting point for the systems design in its barest form. It leaves out many subsidiary factors which were thought to be obvious, such as the fact that mobile equipment must be small and electrically efficient, base station equipment will have to operate unattended on remote hill-top sites and so must be very reliable. It also leaves out other factors which become apparent and are brought into the argument when detailed areas of the system design are considered.

3.2 Interaction Between Subsystems

3.2.1 General

When the design of any piece of equipment was considered, the first requirement was to define its input and output parameters and their magnitudes. The quantification of magnitudes obviously depended on what was delivered by the previous element in the communication chain and what was required by the subsequent one. But treating the system design as a 'green fields' exercise meant that virtually nothing had to be or could be taken as a hard and fast parameter. If some initial values were postulated for certain areas of design, then even if there was even one area of uncertainty this meant that all system values were uncertain.

Since two way communications are required the communications chain is established in two parts. One termed the 'outgoing', 'go', or 'outward' path is that from the control to the mobile; the reciprocal one from mobile to control is the 'return' or 'inward' path. The implied viewpoint is that of the controller or the base station. Design of the system is complicated by the fact that a number of design areas are common to both paths such as: a common piece of equipment in the mobile (the transmitter-receiver), common mobile aerial, common base stations and towers, etc. which means that each path cannot be considered entirely separately.

The return path is the least complicated of the two. Communications are required from one mobile (at a time) to one control, thus there is only one transmitter in the mobile segment. Whilst there may be a number of receiving base stations each receives only the one transmission and then relays this onto control. The usual practice is for each return link to operate on a dedicated frequency so that there will be one link receiver at control for each base station (for each channel). If more than one base station receives the mobile's transmissions then all the various received versions can be presented to control at audio frequencies. This makes the combining or selection of the received signals a relatively easy matter and often a 'best signal selector' is used which may be presented with information via the return path which is a measure of the carrier strength at the base station receiver.

The outgoing path is inherently more difficult. Communications are required from one controller to either one mobile, a group of mobiles or even a whole fleet. Even for the single mobile case its location is likely to be unknown and it is by no means certain that any one base station will be the best or even acceptable, for the duration of transmission. Thus some means will be needed to either; select the optimum base station, continually compare its performance with the potential of others, and re-route the transmission if desired; or to allow the mobile to select or combine the transmissions at radio frequency. For this reason, and the fact that it was the outgoing path which has had to be relocated in frequency, attention has concentrated on the outgoing path.

A diagram was constructed for the critical outgoing path to show how the elements of the system interfaced with and influenced each other. This is shown in Figure 3.2.1. It has something of the form of a flow diagram but is probably more correctly termed an interaction diagram.

3.2.2 The Interaction Diagram (Figure 3.2.1)

The interaction diagram is presented in a form which should be applicable to a number of mobile radio systems of differing requirements. For this general requirement five elements were considered to be 'inputs' to the system design and immutable.

These were:

- Grade of service
- Frequency band(s) assigned for mobile operation
- Area to be served
- Number of control rooms per scheme and their locations
- Frequency bands assigned for fixed link operation.

It will be seen that the diagram has no 'output'. This is a consequence of the fact that all elements were variables and the output of one element affected the input of another. This perhaps needs explaining by an example. The output of the mobile receiver might be thought of as being a true output if it could be defined in terms of say level, noise, distortion etc. But such a given performance could be obtained for a range of say, receiver input sensitivities/base station transmitter powers. Thus an output, in the sense of a hard and fast specification of receiver interfaces cannot be decided even for this simple situation. An output from the diagram can only be obtained when all the loops shown on the diagram close. By this is meant that, when the output of design parameters or interface performance of an element at the bottom of a return loop line are fed back as input parameters to the element at the top of the loop line, the first mentioned output parameters are not changed. ie. the loop is static or closed.

When all loops are closed then all interface parameters are defined and the detailed design of the equipment(s) within each element can be started. If, following the detail design, it is found that the interface parameters cannot be met then there is a certain reaction to the element adjacent to this in the chain and possibly to many more or even all the elements. This is a manifestation of the need for the system designer to have, or be able to obtain, knowledge of equipment details - at the very least to the level where he can question assumptions of performance and if necessary perform experiments and evaluations to substantiate the assumptions.

3.2.3 Discussion on the Interaction Diagram

Since a little time will be spent discussing aspects of the diagram it is worth meeting some possible criticisms. Whilst I doubt that fault would be found with its principle, it could well be argued that it was a) over simple in that more paths and connections should be shown, or b) covered the elements in the wrong order. The first of these points is probably true but it is thought that the main paths are shown and that others would only serve to unnecessarily complicate the issue at this stage. In reply to the second point one can only draw the analogy of the chicken and the egg. Since the elements are in cyclic/circular order it is perhaps not important in what order they occur, however, the order chosen is the result of much thought and therefore to be defended. Again an example is of use.

It could well be argued that the 'propagation' element is wrongly located at the tail of the mobile branch, and that it should precede the 'number of base transmitting sites' (located at the top of the main trunk of the diagram) since they will be determined by it and the three input elements at the top. Given a purely theoretical exercise this would be true but in virtually all the practical cases sufficient engineering feel (another name for experience?) should be in the designers hands to be able to make a good guess at the number of sites required and the real question will be what existing base stations can be used. Only after finding that the design loops will not close would serious consideration be given to the requisition of new sites due to the extreme difficulties of developing them. Therefore the propagation factors are placed as shown in the sequence of traversing the diagram more as a check on the first guess.

3.3 Use of the Interaction Diagram (Figure 3.2.1)

3.3.1 The Trunk

A traverse of the interaction diagram will show some of its uses and properties.

It will first be necessary to have defined the input parameters of:

1. grade of service

2. mobile frequency band

3. area to be covered

4. number of controls, and

5. fixed link frequencies

For the case under study here these could all be defined; (grade of service and area as already described, number of control rooms decided by the particular county scheme, and the desire not to change link frequencies) except that of the new band of operation. This was to result from negotiations since at first sight there were no suitable bands available - they were obviously all in use. However the indications were that even if operation could continue in the VHF region then frequencies significantly higher than the old 100 MHz band would have to be used. This would make it very difficult to pair with the unthreatened 80 MHz band - the separation between would be too large. Therefore both the outgoing and return bands might well have to be relocated. Whilst it was envisaged that the search for one replacement 4 MHz band would be difficult, that for two, with a suitable spacing, would be nigh impossible. In fact the frequency band element for this exercise was not an input but a variable which had to be incorporated in the considerations. So, in order to allow this preliminary excursion, the assumption will be made that a band or bands will be made available in the VHF region but at frequencies higher than the ones to be vacated.

Whilst the interaction diagram is thought to be applicable to the general land mobile radio situation, the input parameters and background information quickly make any traverse specific to one particular type of system; therefore discussion will be limited to the situation under study here.

From the input parameters (however ill-defined in this case) attention passes to the number of base transmitting sites. Since the expected change of frequency is not too large, then simple theory predicts no significant difference in the level of carrier received in a mobile compared to the old bands. Therefore the number of sites is approximately the same as before and the same locations can be used. One of the significant factors which could affect the signal levels is that attenuation of radio signals through trees increases with increasing frequency and diffraction is less advantageous so it should be borne in mind that there may well be a need for some enhancement by higher transmit powers/extra base stations/etc. On the other hand the level of man-made noise (ignition etc) decreases with increasing frequency.

The next stage reached is 'area coverage/modulation'. This is a thorny one and an initial survey of literature for the area coverage situation shows that, whilst there have been a number of contributions to the subject, a marked degree of commercial interest can be detected and there are no firm comparisons. The modulation question has been a contentious issue in the field of mobile radio for a very long time, again with no firm unbiased comparisons - certainly for the situation which is likely to pertain here ie. 12.5 kHz channelling. It is however obvious that modulation and area coverage are interlinked. It will be necessary at this stage, therefore, to highlight this as an area which needs urgent and extensive attention. To enable the traverse to continue an initial assumption will be made that again the 'old' system will be used. This is quasi-synchronous AM - but even if this were to be adopted as the final choice much work would be necessary to put it on a firm design basis.

3.3.2 Fixed Link Branch

A choice of paths can now be made. The easiest appears to be the left-hand one through the fixed link path. Since this is postulated as not changing then all the factors can be assumed to be known and compatible with each other, so that the complete loop through the fixed links and back through the first three elements can be made, bringing the location back to the same choice junction. Several points concerning this loop are worth making at this stage none the less.

The transmitter/receiver design loop element encompasses all the factors shown in greater detail for the base station-mobile transmitter/receiver design loop enclosed by the box shown. Since the mobile and fixed link elements occupy the same base station and in particular their aerials share the same tower, there will be some cross linking for the aspects of transmitter/receiver compatibility (mobile to fixed and vice versa), aerial and feeder design, and the inter-modulation aspects. Since the fixed links are already higher in frequency than the mobile services, the expected move could bring the two closer in frequency or possibly to a disadvantageous relationship (eg. separated by a receiver's intermediate frequency). So a note of caution is sounded here against thinking that the fixed path could be ignored in future traverses. There is also the constraint of the type of area coverage system - yet to be determined - which must not be ignored.

3.3.3 Mobile Branch

Back at the three way choice junction again. Taking the mobile path this time it is reasonable to assume that a suitable receiver can be manufactured for any of the postulated frequency bands (although its cost may well depend on the actual band) and that the performance, to a first order, will be no worse than those in present use - in fact, being more modern, improvements in technology may well give improved performance. But another note of great caution must be made here, for if the modulation/area coverage situation is unusual or particularly if a new modulation system and/or narrower channelling were to be used, then these assumptions would be far from valid.

The field strength required at the mobile can readily be calculated given the receiver performance, frequency, and aerial type. As has already been said simple calculations show that transmitter/receiver performance including propagation and aerial factors is unlikely to be very different to that in the old 100 MHz bands. It is worth noting at this stage however that the field strength required even on this simple assumption is not independent of frequency and this will couple with the other non negligable propagation factors of noise and interference.

Having passed through the propagation element we return again to the three way junction via the first three elements.

3.3.4 Base Station Branch

The first element in the base station path is that of radiated powers. It has previously been argued that powers higher than those in current use may well be necessary. This is unlikely to pose a significant problem for the next three elements of the chain but the fourth one of the base-receivers' compatibility will be affected. If the separation between go and return frequencies is greater than in the old system then this aspect should not present a difficulty. If however the base receive band has to move for the previously mentioned reasons (section 3.1.1) then a closer separation (both in absolute frequency terms and more importantly in fractional terms) might well result and then close attention will have to be paid to this area.

The return loop for the base station takes us first back to the radiated powers because there may well have been a discontinuity detected between the base receiver performance and the impinging transmitter radiation which requires a reassessment of either the radiated powers, or the subsequent base station path elements (including the receiver chain) to achieve a degree of compatibility.

But, before finishing this path, attention must be paid to the intermodulation aspects. Unwitting and possibly unsuspected non-linearities of the transmitters, aerials and tower structure, and the mobile itself will cause a mixing of the transmitted signals and these could affect either the normal operation of the mobile, or the base station receivers, or other users of the spectrum. Again a closer spacing of transmit and receive frequency bands at a base station could place this aspect in great significance as could operation in a bandwidth narrower than that used previously.

Having considered this aspect then the sequence can return to the start through the radiated powers element again, and if necessary through the base station design loop.

3.3.5 Summary

Thus a complete traverse has been made of the interaction diagram to show that at least all the major system factors have been considered and a degree of their interplay demonstrated.

Even this preliminary and cursory traverse has highlighted some areas which require detailed and urgent consideration and, whilst not wishing to be sanguine about any aspect of the systems design until the loops have been closed; other aspects can be given less attention at this stage.

The areas of prime interest could be stated as:

a) What frequency bands to use. This should really be an input parameter but in this real situation was open for evaluation and discussion.

b) The area coverage/modulation question.

c) The base station radiated powers.

d) The base transmitter/receiver compatibility.

e) The intermodulation question.

f) The impact of the fixed links on the area coverage requirement.

g) The impact on the fixed links of the new transmit (and possibly receive) bands

h) Propagation factors in the new bands.

These areas of study therefore became my priority tasks and the work on them will be described in the following chapters.

4. Propagation Factors I

4.1 Basics

The propagation elements are shown on the interaction diagram (Figure 3.2.1) as coming into force only after a number of other elements have been considered. This assumes that the underlying or basic factors pertaining to mobile operation are already understood to the mobile radio designer. These factors will be covered here in very skeleton form, partially to set the scene for subsequent chapters, and partially to highlight some particular aspects. The remaining parts of the heading will be dealt with later, in the chapters on Propagation Factors II (chapter 6) and Propogation Factors III (chapter 8).

The very basic view of propagation for VHF mobile radio is depicted in Figure 4.1.1. It shows a mobile receiving a transmission from an elevated transmitting aerial at a base station. The received signal consists of the combination of that received directly (free space path) and that which has been reflected from the ground. The ground is assumed, to a good approximation, to be a perfect reflector and for vertical polarisation there will be an inversion of the vertical field vector. Thus the two waves tend to be in antiphase and the degree to which they cancel is determined by the phase angle difference of their field components, and it is, in turn, dependent on the frequency and difference in path lengths.

The field strength at the mobile can readily be derived (see Matthews 1965) as

$$E_m = \frac{k_1 P^{\frac{1}{2}}}{\lambda} \cdot \frac{h_1 h_2}{r^2}$$

where h_1 and h_2 are the heights above ground of the transmitting and receiving aerials respectively, r is the distance of the mobile from transmitter, P is the radiated transmitter power and λ the wavelength of operation. K_1 is a constant of proportionality dependant on the units used. On this model the field strength, all other factors being equal, is proportional to frequency.

The mobile aerial is usually a $\frac{\lambda}{4}$ whip mounted on the vehicle roof (or, with inferior performance, on the wing etc.) The effective length of such an aerial is proportional to its length which in turn is proportional to the wavelength of operation. Hence the voltage received by the receiver will be

$$E_r = \frac{K_1 P^{\frac{1}{2}}}{\lambda} \cdot \frac{h_1 h_2}{r^2} \cdot K_2 \lambda = K_3 P^{\frac{1}{2}} \left(\frac{h_1 h_2}{r^2} \right)$$

and the received signal power is

$$P_r = K_4 P \frac{h_1^2 h_2^2}{r^4}$$

where the K_n's are constants of proportionality.

On this simple model several significant features are apparent:

i) receiver power is an inverse 4th power function of range rather than the inverse square law of the free space case.

ii) there are great benefits in having aerials as high as possible.

iii) received power is independent of frequency.

The last factor assumes that the type of aerial at the base station and that at the mobile are independent of frequency eg. dipole - quarter wavelength whip, 3 element yagi - 5dB gain colinear etc.

The very simple model used above is usually extended by considering firstly earth curvature and, secondly by making some allowance for the fact that the transmit aerial tower is itself usually located on a hill-top, so that the aerial height to be used is dependent on the range and details of the topography of the situation. The topographical model is then extended to include non line-of-sight paths which may be obscured by one or more topographical features (eg. hills). The final stage is to show that even if all these factors could be accounted for rigorously then there is still a high degree of variability caused by multipath.

The multipath arises due to the reflection of the transmitted signal by objects ranging from distant hills to nearby structures such as buildings or anything larger than some fraction (say one tenth) of a wavelength. This situation has been analysed by Clarke (1968) on the assumption that a spectrum of multipath signals would be received with random arrival directions, magnitudes and relative phases. The predicted Rayleigh probability distribution (see appendix D) for the received signal strength resulting from this has been confirmed in a number of measurements. The statistical measure of the distribution is however constant over the shadowing effect of hills, trees and buildings; so that two statistical distributions are needed to cover the variability at any given range from the transmitter. These are a log normal distribution and Rayleigh in combination, with the former having a standard deviation of some 6 to 12 dB (French 1979).

A number of propagation models have been proposed in order to predict the received level; that of the CCIR (1978) can be taken as acceptable at least for initial predictions. Following extensive measurement exercises there have been a number of proposals by several authors for values of the range exponent different to 4, but it has been shown (Parsons, Ibrahim, and Samuel 1980) that when the extreme variability of the measurements are taken into account the difference in fit of the experimental points for either the proposed law or the fourth power one is small. Therefore this value of 4 will be used here at least for the purposes of comparing re-use strategies, area coverage and modulation.

4.2 Measurements

The Directorate has had a practice of measuring the levels received at a mobile from a base transmitter before implementing a new base station site. The measured values are plotted as colours representing bands of signal strength along the roads traversed. This is of use to the site user in assessing its acceptability. The measuring system has been automated for use in assessing coverage for the new bands and the measured information (averaged over several metres) is available for machine analysis and plotting. Suitable software has been developed to analyse the data so that plots of signal strength against range can be made. Figures 4.... to 4.... show some of these which were made once the frequency bands were known. They show that the fourth power assumption is valid and also that there is a very significant difference between the ranges expected to be achieved for the 10% and 90% cases compared to the 50% (median). Details of the measuring system and discussion of the accuracy and validity of the results is discussed in Appendix A.

The 10% curve is of interest when considering realistic frequency re-use distances, and the 90% one is that which applies to the desired grade of service.

The curves show the signal received by a mobile radio receiver connected to a quarter wavelength roof mounted whip aerial when the transmitter delivers 10 watts to a dipole mounted on a 30m mast. Giving the results in this fashion - although not normalised - avoids possible uncertainties on the part of the survey team and the analysers/designers. These uncertainties usually stem from the use of the term 'path loss'; it seems to be a moot point whether this is between isotropic aerials, dipoles or even field strengths. There is then the need to consider the performance of the vehicle aerial and whether its nominal 37.5 ohm impedance has been matched to either the vehicle 50 ohm feeder and/or the nominal 50 ohm input impedance of the receiver.

From these graphs average figures can be derived for K_4, (see Appendix A) these are for the median $\qquad K_{50} = 4 \times 10^{10}$

for the 90% exceeded $\qquad K_{90} = 4 \times 10^{11}$

where d is in km

It is often desirable to give a formula readily usable in meaningful units and scaled in logarithmic terms such formulae can be stated as:-

$$P_{50} = P - 84.2 - 40 \log d$$

$$P_{90} = P - 94.2 - 40 \log d$$

Where P_{90} and P_{50} are the received powers for the 90% and 50% cases respectively and they and P are expressed in dBW (dB relative to 1 watt) and the range d is expressed in km. These formulae do not include h_1 and h_2 since representative values for these have been incorporated by the process of averaging many plots taken on a range of terrains in England and Wales.

It is of interest to note the ranges predicted for the two cases if typical values are inserted for effective radiated power (from a dipole) of 50W (+ 17dBW) and the minimum desirable power for a typical receiver of -131dBW (4μV emf)
(4 Vemf). These are:-

50% of location covered for a range of 39 km (24.5 miles)

90% of location covered for a range of 22 km (13.8 miles)

The predicted areas covered are of course circular since omnidirectional aerials have been assumed. In practice some use might be made of directional types at the base stations and in any event the survey would be conducted for each area to show the actual coverage. At this stage of the argument however the assumption of circular coverage areas will continue.

It is probably worth commenting at this juncture on the use of horizontal polarization. It is often stated that this cannot be used for mobiles since the horizontal field strength falls to zero at the ground even for 'free space' paths only. Since the predicted field strength for vertical polarization on the basis of a sensible model has been shown to also fall to zero at ground level this argument is not valid. It is somewhat more understandable if the argument is modified to the fact that the horizontal field strength will be near zero on the (conducting) roof of a vehicle. The main reason however is probably more the difficulty of achieving omnidirectional coverage from a simple aerial at the base station and even greater difficulty at the mobile.

5. Area Coverage

5.1 General

One of the requirements was to give the users at least one channel that would be available anywhere in his province - for the typical county this was equated to covering an area of some 50 x 50 km (30 x 30 miles).

From chapter 4 the range of coverage for even a 50W transmitter using a typical hill top location was shown to be only some 13.8 miles for a reasonable grade of service. So even if one transmitter could be centrally located in the county then the area covered would be insufficient to say nothing about the coverage of the ill defined shape of a real county. Therefore more than one transmitting location will be needed per county - in fact an average of four is a reasonable prediction. It will be recognised that whilst coverage might be thought to be greater in flat areas, the fact that high sites are usually hard to come by in these areas counters this. In hilly areas on the other hand a good elevation for the base station can easily be achieved but the likelihood of shadowing by the hills is also greater. (It follows that the easiest area to cover is a low-lying one with peripheral hills such as a coastal plain with inland mountains.)

The question to be addressed therefore, becomes one of how the system should be operated so that the several base stations give the service to a mobile in an efficient manner. Both directions of working are to be considered but for the reasons discussed previously the outgoing path from base station to mobile is the difficult one. The term efficiency here applies in part to the conservation of radiated (or indeed prime) power but more pertinently to the aspects of:- bandwidth used, reuse distance, and time to transmit the message. All these can be covered by the expression spectrum efficiency.

Three possible modes of operation can be envisaged:-

i. base stations on a common scheme share a channel but only one is excited at a time.

ii. each base station assigned its own channel, and can be excited independently of any other in the scheme.

iii. base stations share a channel and more than one can be excited at one time.

This is a choice very similar to that faced when wishing to combine or multiplex signals onto a common highway where the first two are termed

i. time division multiplexing

ii. frequency division multiplexing.

The third has no direct analogue in this respect (except for a form called spread spectrum modulation which will be dealt with as a subset of the modulation topic), but from the fact that simultaneous transmissions are permitted it is sometimes referred to as 'simulcast' (simultaneous broadcasting) other terms are common frequency, common channel, synchronous and quasi-synchronous. Common channel is probably the best term, with spaced carrier, synchronous, and quasi-synchronous as subdivisions.

The techniques to be examined are then Area coverage by:-

i. time division operation

ii. frequency division operation

iii. common channel

and these will be discussed in turn.

5.2 Time Division Operation

5.2.1 Considerations

The situation to be considered is that shown in Figure 5.2.1. The total area to be serviced is covered by a number (say four) basestations, each radiates on the same frequency F1 but only one is activated at a time (Tn). It is assumed that no part of the service area is inaccessible from at least one base station and there may well be extensive overlaps in coverage from two or more transmitters.

For the outgoing path the message will be assumed to have been originated at the control so that the control desires to pass a message to one mobile. If the location of the mobile is known then only the transmitter serving that area needs to be energised, and the message can then be passed. In general however the location will not be so well known for the type of mobile which needs this county coverage mode of operation so that the transmitters will need to be activated in sequence in order to ensure that the correct one will be in operation at least once. Certainly all the base stations need to have transmitted the message if either:- it is desired to communicate with the whole fleet of mobiles, or it is desired that all mobiles should be aware of the traffic on the channel. The user frequently regards this as a valuable feature of area coverage as a means of keeping senior officers, who are mobile, in touch with the overall situation and enabling all users of the channel to have background general awareness, and instant awareness, of a developing situation so that they can use their initiative in supporting the operation of others.

An intermediate mode of operation can be proposed where the control quickly cycles around the base stations with a short message with the purpose of eliciting a response from one desired mobile before transmitting the main, assumed larger message, from one selected base station. The problem with this approach is the time which should be allowed for the mobile to respond. This might reasonably be set at say 8 seconds, and the query (such as "alpha-romeo 3 receiving?") will take another 3 seconds, so that the total per base station is 11 seconds. For a four station scheme the total could therefore be some threequarters of a minute - just to enquire whether a mobile is in a position to respond to a message. Obviously if the subsequent message is a very long one then this might be viable but if it is short then it is certainly not. The cross-over point for various response times and message lengths is shown in figure 5.2.2. The assumptions made here are that the query time Q is 3 seconds, the time to address a mobile A is 2 seconds (eg "alpha-romeo 3") and that the mobile will in general respond after an average of two queries. So with the time waited for a respone R seconds and the message length M seconds the query sequence will take 2(Q+R) + M seconds, and the sequential method 4(A+M) so that the query method will be shorter if

$$4(A+M) > 2(Q+R) + M$$

Figure 5.2.2 shows the boundary for message length as a function of time to respond for the values of A and Q of 2 and 3 seconds respectively. From this it will be seen that messages longer than 6 seconds should be handled by the query method for the postulated response time of 8 seconds.

Whether the sequential method or the query one is used however, another factor comes into play. This is the blocking of the return channel. Whilst the channel is assigned to outgoing transmissions incoming ones are unwanted and should not disturb the system. Therefore the engaged 'pip-tone' should be radiated from the transmitter not being excited with the query or message. This is necessary in order to inform mobiles who cannot hear the main transmission that the channel is in use and that they should not transmit. The transmission of the pip-tone in these circumstances raises two problems. The first is that, in areas of service overlap where two pip-tone transmissions are received, then there will be unwanted effects on all mobiles (all mobiles maintain a listening watch). The effects will be an audio beat note at the frequency of the difference in transmitter carrier frequencies (if the carriers are not accurately controlled then this could easily be around 1 kHz, effectively right in the middle of the audio band and potentially very disturbing), and distortion of the pip-tone if the modulations are not matched (see the section on modulation for fuller details). The second problem here arises in those areas where transmissions of the pip-tone overlap with those of the query/message. Here the beat note will be present again (with its strength greatest where the two received signals are equal) but there will be severe conflict of the two modulations. It might be argued that the so called capture effect of frequency modulation could be of benefit here but the effect is minimal with the deviations which can be used in a 12.5 kHz system, and in any case the advantage which capture may confer for carriers differing in amplitude by a ratio higher than the capture ratio, is more than offset by the fact that when the carriers are closer in amplitude than this ratio the result is a total corruption of both signals (see sections 6.2 and 7.4.5).

The time division approach obviously does not make 100% effective use of the time available. Now emergency service communications need to be scaled to meet virtually peak demand. This is in contrast to the design of telephone systems or even mobile radio telephone systems or traditional private land mobile radio systems; for there it is considered uneconomical to approach peak demand requirements due to the under-use of resources at other times.

During peak demand for the emergency services the requirement is such that the controller is fully active for long periods. Measurements have been made on existing systems operating in a common channel mode (quasi-synchronous) which show near 100% occupancy of the channel for periods longer than an hour. To cater for this loading on a time division system with say a 50% time efficiency (ie. messages occupying the channel for only 50% of the time) would mean that not only were two channels required to be available but also two controllers. Thus the time inefficiencies of a time division system are effectively translated into frequency inefficiencies (more channels required) and resource inefficiencies (more manpower). If the time inefficiency is compensated for by more channels then this immediately conflicts with the specification for area coverage which implies that all mobiles should receive all the messages on the channel. This can no longer be achieved since the message channel has been split to cover more than one frequency channel at the same time.

5.2.2 Summary

Summarising therefore the time division system may cause the mobile annoyance by virtue of receiving two carriers at once and be ineffective by corrupting transmission; these effects will occur in areas of overlap of service areas for two or more base stations.

The inherent time inefficiency of the technique will, in order to achieve the service necessary, be translated into spectral inefficiencies and result in the aims of the area coverage requirement not being met.

This technique could therefore only be adopted if other possibilities showed even poorer performance.

5.3 Frequency Division Operation

5.3.1 Considerations

Figure 5.3.1 shows how the transmitters and frequencies would be arranged, it follows the form of figure 5.1.1 since it is intended to cover the same area from the same base station locations. In this case however each base station operates on its own frequency (F_1, F_2, F_3, F_4) and they can all transmit at the same time (T1). Providing each mobile is tuned to the correct frequency in each coverage area, then a control originated message can be simultaneously transmitted from all base stations in the scheme and received by all mobiles. The proviso in the last sentence is not one to be passed over lightly however.

As has been pointed out before, the aim of the area coverage channel(s) is that the mobile user can operate as if he were tuned to just one RF channel. There can be no reliance placed on the mobile operator changing channels as he goes from one service area over to another, any changing would have to be done automatically. One method of achieving this would be to have an automatic vehicle location (AVL) system. Knowing the mobile's geographical position then the correct base station for that area can be deduced and the mobile receiver instructed to tune to that frequency. This requires several important elements in the 'tuning' chain. Firstly the mobile receiver must be capable of channel changing under external control, ie be 'programmable'; this is not a problem particularly if the receiver's local oscillator is derived from a frequency synthesiser. Secondly the source of the location information needs consideration; if the AVL processing is performed in the mobile then the information is easily available (even if more processing is needed to put it in 'base station service area' rather than simple geographical location form) and can be used directly. It is more likely that the actual location processing would be performed centrally (ie at control) and then there is the problem of communicating this to the mobile. This leads to a circular argument in which the information can be passed to the mobile over a radio channel only if it already knows which channel it should be tuned to. Even if this were resolved the last element of the chain requires that there be an installed viable AVL system. These do not exist in the emergency services as yet and, despite considerable efforts on the part of the Directorate to promote them, and on the part of industry to provide them, there is no system on the horizon which could be considered to be capable of implementation by the time needed to meet the WARC programme.

An alternative form of AVL which could be used here is one which only senses the base station service area most appropriate to the mobile. This provides only the crudest of location information in the geographical sense but is all that is needed for this purpose. It would operate by setting the mobile radio receiver to search and find which of the base station frequencies was the strongest and latch onto that. Scanning receivers exist which would cycle through the few (four in the case we are considering) channels of one area coverage scheme many times per second. Indeed such a receiver has been proposed commercially for the emergency services. Continuous base station transmissions would not be needed for the operation of this technique if the scanning and latching rate of the mobile is fast enough to acquire a suitable channel in the time between the base station's carrier becoming actuated and the start of a speech message. This should be possible. Moreover the cost of such a receiver whilst being higher than would otherwise be the case should not be significantly so due to the low cost of digital processing and the fact that the receiver would almost certainly have a synthesised local oscillator system, which is inherently frequency agile.

There is in any case a need to recognise the fact that a facility such as area coverage, or indeed that of having more channels in a given band of frequencies, entails at least some element of penalty. This penalty may be manifest as either greater cost, poorer performance, or of course a combination of the two.

There will be no problems of beat notes in the areas of base station overlap but in these areas neither will there be any degree of reinforcement of one weak signal with another weak one. The fact that say four frequencies have to be used to provide just one channel would appear to give inefficient use of the spectrum. This is not necessarily the case however, especially if it is compared to the case of one high powered single frequency equivalent. The section on frequency re-use covers this aspect in more detail.

5.3.2 Summary

Frequency division operation should be viable in that it should be possible to realise the equipment but this will incur a cost penalty. The degree of spectrum efficiency will be covered in the section on frequency re-use.

5.4 Common Channel Operation

5.4.1 Considerations

The symbolic diagram of operation is shown in Figure 5.4.1, the same channel (F1) is used in all service areas at the same time (T1).

In those locations where a mobile receives transmissions from one base station only then, assuming that this is of sufficient level, the system will work satisfactorily with no restrictions on the base station equipment or the mobile equipment compared to the single base transmitter condition. Problems arise however where the mobile is affected by two or more base stations transmissions ie in the overlap area. Here there will be interference (a diffraction pattern at radio frequency) between the signals and the effect will depend on the nature and severity of the interference, the type of modulation used, and the performance/design of the receiver. The details of operation in this mode are considered in the chapter on modulation but a simplified approach will demonstrate the basic factors.

Consider firstly unmodulated carriers to be radiated, then when say two of these are received by a mobile, their combined effect will have the appearance of a single sideband modulated carrier (SSB) with the stronger carrier modulated by the weaker at a frequency corresponding to their difference frequency, as shown infigure 5.4.2. Thus there will be imposed on the stronger carrier both amplitude and phase/frequency variations. An AM receiver will respond to the amplitude component and an FM one to the frequency component. Considering now modulated carriers, providing the modulation spectra of the transmissions do not overlap in frequency then a linear demodulation system will avoid interaction between the transmissions and the receiver will present the sum of the modulation components plus a beat note at the carrier difference frequency.

SSB is the only linear demodulation system considered and even here there will be doubt as to which of the components is the carrier and which the modulation. The next nearest linear system is probably AM, and the common channel technique has been in use in the Directorate's systems for many years using this technique. It requires however a separation between the carriers greater than twice the baseband bandwidth to avoid spectral overlap. A figure of 6 kHz separation has been used for this, producing an audio beat note at the demodulator output of 6 kHz which can be filtered easily from the top baseband frequency of 3 kHz by a low pass filter. The spectral distribution for three transmitters is shown in Figure 5.4.3.

The system has served well for the period when the channel bandwidth necessary to accommodate the spaced carriers could be tolerated.

If FM were to be used on the same basis then the allowable deviation would be very small in order to keep all the modulation sidebands from one transmitter within the $\pm$ 3 kHz used by the AM system and avoid overlap. Such small deviation would render the performance under even single station reception conditions worse than AM since no use would be made of the ability of FM to exchange performance for bandwidth (deviation). Spaced carrier FM systems have not been implemented.

Spread spectrum (SS) modulation would seem to be a possible mode here but in order for the instantaneous spectra of the transmitters not to overlap (most easily perceived in the case of a frequency hopping system) the transmitters would either have to have different code sequences which would negate the idea of receivers all being 'tuned' to the same channel and falling to the same considerations as for the time division situation, or would have to transmit the same codes against differrent starting time references. How a receiver would behave in this latter situation where it might be required to advance or retard its acquisition system as rapidly as changes were to occur from one transmitter to another is an unknown. Thus spread spectrum appears non viable in this situation.

Spread spectrum signals would require considerably more bandwidth than even the 25 kHz previously available but they cannot be ruled out simply on this ground, since their energy density (watts/Hz) is low which would enable the same spectrum to be re-used on a very close spacing - indeed they can share the same spectrum to a moderate degree (determined by a mutual interference basis) from one transmission site. The other forms of modulation considered are however more energy intensive and, for reasons given earlier, channel bandwidths no greater than 12.5 kHz can really be contemplated. Spaced carrier operation is therefore not admissible. Even with AM limited to an upper audio frequency of 2.5 kHz (not very good quality) only two stations at the most could operate within the receiver's IF passband of 3.75 kHz and these would have to work on a vestigial sideband basis.

The previous statement of non overlapping modulation spectra can be overruled however if the modulation spectra can be considered common to a carrier. This can be seen for the case where the two transmitted carriers are synchronised and their modulations identical, in this situation the mobile would not be able to distinguish between receiving the two transmissions or receiving from one transmitter with a strong reflected component - the normal mobile situation. If the mobile were in motion in the single station with reflection case cited above then, by virtue of the Doppler effect, the carrier (and modulation) frequencies of the two signals would be different. Since the mobile copes with this situation also, then it would seem that the two separate transmissions of an area coverage system do not need to be synchronised, and perhaps the modulation components can be a little mismatched. Thus operation in a quasi synchronous (QS) transmission mode appears feasible.

On this basis AM, FM (and SSB) could all be contenders.

Alternative forms of argument for spaced carrier and quasi synchronous operation, have been advanced which arrive at the same conclusions as that covered above. They hinge around ensuring that the carrier frequency beat note lies either above the audio output spectrum (spaced carrier) or below (QS) so that either they may be filtered out, or they are in a region which can be tolerated aurally. Whilst such arguments have value - and will be brought into discussion in the modulation section - they do not account for the 6 kHz spacing needed for spaced carrier operation, the inabilities of FM to perform in this mode, and the differences in tolerances to modulation alignment required between spaced carrier and QS systems which will be covered later.

Quasi synchronous AM and FM systems have been implemented in the recent past (the former by the Directorate) with varying degrees of success, but with scant consideration to the design parameters and naturally no accounts have been given of the failures. Only one successfuly implementation is however necessary to show that the principle is sound - success may depend on the details. Such details could be customer tolerance to:- differences in carrier frequencies modulation mismatch, design of receivers etc; these will be covered in the relevant sections of chapter 7.

5.4.2 Summary

Common channel operation would seem to give the coverage required but due to bandwidth restriction will result in operation in a quasi synchronous mode. That both AM and FM systems can seemingly be made to work lends weight to this but the failures of some installed schemes point to investigation being needed to determine the critical design features.

Despite the viability discussed so far, the question of their spectral efficiency/re-use factors have not been addressed. This could apparently be a vulnerable area since transmitters on the operating frequency will be geographically dispersed and located away from the centre of overall coverage.

5.5 Transmitter Power Consideration

5.5.1 General

This chapter on area coverage demonstrated at the beginning that the area of a typical (or for that matter any practical) county cound not be covered from one base station; four were postulated and the subsequent arguments were based on that as being a typical number. Other very important advantages accrue from the increase in base station numbers; one of the foremost of which is the attendant reduction in transmitter power needed - both for the base station and the mobile. It is important for the latter due to the associated qualities of, prime power, size, heat dissipation, interference to other services etc. For the former a reduction in transmitter power is not only good engineering but it will be shown in the chapter on intermodulation that any reduction is highly beneficial in reducing interference to that cause and that intermodulation effects pose a major constraint on system design. Thus minimising base transmit powers will be a major target.

5.5.2 Four Base Stations

To illustrate the nature and actual magnitudes of reductions which can be obtained consider the artificial case of a county of square shape and take the single transmitter as a reference. Figure 5.5.1 shows a plan view of the county covered by the one transmitter. The range of coverage of this is r_1 and the area covered on the very simple assumption of the whole county being tesselated by similar squares is $A_1 = \frac{r_1^2}{2}$. Figure 5.5.2 shows the case where four transmitters are used each having a range r_4 and covering their own areas $A_4 = \frac{r_4^2}{2}$

Figure 5.5.3 is a cross section of the county along a diagonal and shows the power receivable by a mobile for the two cases. It assumes the inverse fourth power law and that the minimum mobile receiver power P_{min} just obtains at the maximum range. Using the inverse fourth power law and assuming that the transmit aerial heights are all the same then

$$P_{min} = \frac{KP_1}{r_1^4} = \frac{KP_4}{r_4^4}$$

Therefore $$\frac{P_1}{P_4} = \left(\frac{r_1}{r_4}\right)^4$$

and with $r_1 = 2\, r_4$

$$P_1 = 16\, P_4 \qquad \ldots(5.5.1)$$

Each of the four transmitters can therefore radiate powers 12dB lower than that of the reference single base station.

Figure 5.5.3 demonstrates that the single base station would provide a power flux which in most of the area is considerably more than that needed. The four station situation is clearly better in this respect but the figure is a little misleading if interpreted simply since it shows only one of the two linear dimensions. The average power flux over the whole area for the four transmitters is $4P_4$ or $\frac{P_1}{4}$ ie a reduction of the overkill on power by a factor of four.

The situation depicted in Figure 5.5.2 is that pertaining to area coverage by either time division or frequency division since in either case each base station will have to provide coverage to the whole of its area in its own right. For the common channel situation however further improvements can be made.

These stem from the fact that all transmitters will radiate simultaneously and co-operatively, so that in overlap areas their signals will add. Whilst the addition is on a voltage basis the chapter on modulation will show that, due to the cancellation of the carriers when they are in antiphase, the addition is in effect equivalent to a power summation.

Therefore the power at point C will be four times higher than necessary and at B twice as high. This means that the base station can in principle be relocated and its power adjusted to just meet the criteria of the three corners of coverage as shown in Figure 5.5.4.

The relevant equations for this are

at A $\quad Pmin = \dfrac{K P_{4c}}{a^4}$

at B $\quad \dfrac{Pmin}{2} = \dfrac{K P_{4c}}{b^4}$

at C $\quad \dfrac{Pmin}{4} = \dfrac{K P_{4c}}{c^4}$

Where P_{4C} is the transmitter power for each of the 4 (county wide) transmitters in the co-channel case and a, b, c, are as shown in the figure. $(a + c) = r_1 = 2r_4$ (the other two transmitters make an insignificant contribution to the power at B - only one twenty-fifth of the total power before they are themselves moved). The variables in these equations only admit of two of the conditions being met; it is necessary therefore to ensure that the other condition is exceeded. The easily soluble case of A and C together shows that condition at B is not met. Therefore it is necessary to consider the transmitter nearer to B than for that situation, since the transmitter needs to be on the line AC in order to give a service to the fourth corner of the square equivalent to that at B, it must move nearer to C, and it will have a power higher than in the A/C case, so that condition C will be more than just met.

Therefore $Pmin = \dfrac{K P_{4c}}{a^4} = \dfrac{2 K P_{4c}}{b^4}$

giving $\quad b = a(2)^{\frac{1}{4}}$

The geometry of the square gives

$$b^2 = s^2 + a^2 - 2as \cos \theta$$

where $s = \dfrac{r_1}{\sqrt{2}}$ and $\cos\theta = \cos\dfrac{\pi}{4} = \dfrac{1}{\sqrt{2}}$

Therefore $\quad b^2 = \dfrac{r_1^2}{2} + a^2 - a r_1$

substituting for b^2 and regrouping gives

$$a^2(\sqrt{2}-1) + a r_1 - \frac{r_1^2}{2} = 0$$

which yields $a = 0.425\ r_1$

and in turn $P_1 = 30.6\ P_{4C}$

This power is nearly half that of the frequency or time division cases (P_4 in eqn. 5.5.1) and nearly 15dB below the single transmitter reference.

An ideally located base station for the common channel case therefore shows significant reduction in transmitter power compared to the other two contenders for area coverage. In addition it will be demonstrated in the section on re-use that the nearer a transmitter can be located to its edge of cover the better.

Although an idealised and artificial situation has been considered with freedom to locate a base station to a fine degree, the principles will apply in practice and use could be made of some directionality of radiation from a non ideally located site.

5.5.3 Multiple Base Stations

The discussion in the previous section was based on the assumption of four base stations per county; whilst this is a reasonable number to postulate it is worth examining the extension of the principle to other numbers.

Since the total area to be covered will be shared between n base stations then, assuming equal shares, each will cover 1/n of the reference area. Therefore each will have a range of $^1/n$ and the transmitter power will be $^1/n^2$ of the reference one.

This argument assumes that the elemental areas will interlock to cover the plane without gaps. The only shapes that will tesselate in this fashion are regular polygons of 3, 4 and 6 sides - equilateral triangle, square, and regular hexagon. It also tacitly assumes that the original shape (itself an approximation to the reference circular coverage area of one transmitter) is replicated by the n elemental areas. This is true only for the case of triangles and squares where n will be limited to the square of an integer. On this basis, then, only a limited set of base station numbers are admissible. It must be remembered however that the true shape of a county will generally best be approximated by a cluster of elemental shapes of the tesselating variety, the more the better the approximation and therefore the sharing of the true coverage area and shape between the elements can be taken as being on an equal basis. So that, if say the county could be approximated to five equal squares, then the value of n = 5 for n can be used to give an indication of the reduction in transmitter power (25 times) possible compared to the reference single transmitter case. The transmitter power reductions on this basis will not be exact but they serve to give an indication of the gains to be made by increasing the number of base stations and become more accurate for comparison when n is high.

5.6 Frequency Re-use

5.6.1 General Survey

It behoves those who plan frequency assignments on a National scale to make as much use of any segment of the frequency spectrum as possible. To this end they seek to ensure that any frequency associated with a communication channel in use in one part of the country can be used again as many times as possible within the country (and indeed there is much international work to ensure that maximum frequency useage is obtained globally). The Directorate of Telecommunications plans the frequency segments allocated to its users on a nationwide basis. Maximising re-use is said to be spectrally efficient and to conserve spectrum; this conservation is perhaps in an odd sense since the spectrum is never consumed and not in need of renewing - it is always there but may not be available for use. Spectral efficiency is usually defined in terms of the number of users per unit of area per unit of bandwidth - say users per square km per MHz. For the purposes of mobile radio users this can be distilled into its inverse 'the number of channels which need to be assigned on a national basis in order to provide one usable channel at any location in that county', the fewer the better.

To assess various candidate modes of use, it is again assumed that all base stations (and all mobiles) are identical and at first that they have a circular coverage area. The base stations are arranged as close as possible to each other but without leaving unserviced gaps between them. On this basis the base stations could be assigned to the centres of any tesselating shape which approximated a circle, the equilateral triangle, squares, or regular hexagons of the last section thus reappear. It is conventional to take the hexagonal option as most closely approximating the circular coverage area and closest packing of the base stations.

Therefore the county is viewed as being covered by equal sized regular hexagons each with a base station as its centre. The mapping thus obtained is obviously artificial and not realistic in that it takes no account of the many variable features - but as always it serves as a vehicle for the purpose of comparisons. The mapping is shown in Figure 5.6.1, where the cells marked A are assigned the same channel frequency. The concept and geometry of the situation has been regularised by MacDonald (1979). In Figure 5.6.1 the distance between co-channel cell centres u is the re-use distance. Obviously the channel A cannot be re-used with the circle shown.

The re-use distance is determined by the level of signal from A_1 which can be tolerated in the next adjacent A cell - say A_2. The worst situation occurs at the edge of A_2 cell where its wanted signal is weakest and A_1 is strongest. It is the minimum tolerable ratio of the two signals received by a mobile at the edge of A_2 which is considered to define acceptability and hence the re-use range; this ratio is called the co-channel protection ratio Γ. It is also conventional, but not strictly accurate, to examine the mobile received power levels along the line joining A_1, A_2 centres and consider that, again, the service areas are circular with range r.

For a given value of Γ then r can be readily calculated from the inverse fourth power relationship of power versus range and this has been done by several authors (see French 1979, Gosling 1978). Whilst the calculation may be readily accomplished for one interfering base station there are several regions of contention in the calculation these centre on.

i. the contributions of the other surrounding base stations on channel A [schemes for minimising this by effectively siting the base stations at three of the corners of a cell and using directional aerials have been proposed (MacDonald 1979)].

ii. the criterion for measuring the protection ratio, and how this is actually measured (primarily, should it be objective or subjective - particularly relevant to comparing different modulation systems).

iii. the errors likely due to the assumption of a smooth inverse fourth power law (the real signal is subject to fading which is distributed on both a log-normal and a Rayleigh basis. This will be covered in the section on Propagation II and Gosling (1978) has pointed out the errors caused by ignoring this in frequency reuse calculations).

These considerations affect the absolute reuse value which can be used, but within this chapter the object is to compare the various area coverage strategies so the comparison will be made for a range of re-use distances.

5.6.2 Single vs Multiple Base Stations

For this comparison the inverse fourth power law of mobile received power vs range will be dragged into use again. Despite the caveats in the last section this can be justified on the grounds of the fact that the law provides the median for the other statistical distributions described above, and that it is generally tractable and used elsewhere.

Figure 5.6.2 shows the received power intensity (on a logarithmic - dB scale) against range for the cross-section between two co-channel base stations. As before Pmin determines the services ranges r_1 and r_2 served by A_1 and A_2 respectively. The distance d is dependant on the protection ratio Γ. Figure 5.6.3 is drawn for the same service ranges, and the same protection ratio, but this time two base stations are assumed to cover the linear service range of each coverage area (implies four base stations to cover each area). The base station powers have been adjusted to just give the same coverage as for the single one. Due to the more rapid rate of fall-off of the power available outside the service range the distance d, and hence the re-use distance u, is less.

These distances can be calculated for the two cases shown.

Let P_1 and P_4 be the transmitter powers for the single and four station cases respectively, and for the purpose of differentiating between the two situations the distance variable will be taken as x and y respectively.

then for the single station case:-

the power received at a mobile $P_r = \dfrac{KP_1}{x^4}$

where K is a proportionality constant

and at the edge of range P_r = Pmin and x = r so that $KP_1 = x^4 P_{min}$

When the range is r + d the power has fallen to $\frac{P_{min}}{\Gamma}$

so that $$\frac{P_{min}}{\Gamma} = \frac{r^4 P_{min}}{(r+d_1)^4}$$

which yields $r + d_1 = r.\Gamma^{\frac{1}{4}}$

Hence :-

$d_1 = r(\Gamma^{\frac{1}{4}}-1)$ and re-use distance $u_1 = r(\Gamma^{\frac{1}{4}}+1)$.

Now, for the four station case

$$P_r = \frac{KP_4}{y^4}$$

and P_r = Pmin at $y = \frac{r}{2}$

so that $$KP_4 = \frac{r^4 P_{min}}{16}$$

and the co-channel level is reach at $y = \frac{r}{2} + d_4$

so that $$\frac{P_{min}}{\Gamma} = \frac{r^4 P_{min}}{16(\frac{r}{2}+d_4)^4}$$

which yields $d_4 = \frac{r}{2}\left(\Gamma^{\frac{1}{4}}-1\right)$

and re-use distance $u_4 = \frac{r}{2}\left(\Gamma^{\frac{1}{4}}+3\right)$

To give this some kind of perspective, a value for protection ratio of 20 dB can be taken (as in used in figures 5.6.2 and 5.6.3).

Then $u_1 = 4.16r$ and $u_4 = 3.08r$

Now the number of channels needed in a re-use area approximates to the square of the re-use distance (since it is an area function) therefore the number of channels required for the four station case is nearly halved!

It will be seen that the separation distance d between the service areas was halved in the case of the four station case. That this should be so - independent of Γ - can be easily seen if figure 5.6.2 is contracted by a factor of 2 in the distance scale only, and then overlaid on figure 5.6.3. It will be found to fit exactly for all situations between the nearest base stations. This approach can be extended to the case of 3 base stations on the cross section line (9 base stations in the area) yielding a separation distance of one third that of the single one. The general rule could be made that

$$d_n = \frac{d_1}{\sqrt{n}}$$

where n is the number of base stations used for a service area. (The assumption that the n base stations can share the coverage is again made). Thus the gap between co-channel service areas can in principle be made arbitrarily small by using a sufficiently large number of base stations!

The physical interpretation of this in the limit of an infinite number of base stations is that each would provide an infinitesmally small coverage area just giving the minimum power, and thus the rate of fall-off of field strength with distance outside the coverage area would be very high. (To analyse this rigourously the contributions of all other base stations in the area would have to be taken into account, in practice only those very close could contribute anywhere near significantly and are unlikely to alter the substance of the argument. This is confirmed in appendix C.). The power diagram would then appear as in Figure 5.6.4.

This sort of technique could be called the 'lamp-post base station' approach. Apart from the cost of the equipment per lamp-post (both receive and transmit - but of very low performance in each case) there is the question of how the signals should reach the lamp-post and be returned from it. Even the outgoing path poses great problems; the initial reaction is to call for a fixed radio link. This introduces extra cost by virtue of the link receivers but suffers from an overriding consideration. To reach all the lamp-posts in the area would require an omni directional transmission from a central control. This is right back to the starting point!!

The total cost aspects should not be neglected either - they are bound to be greater by virtue of extra numbers even if these are of low power in the case of the main transmitters. The system could however be realised in principle using either each lamp-post in an on-frequency - repeater mode (the stability problems would be immense) or by using land lines. This then becomes a "base station on a telegraph pole" scheme; and could be thought of as one form of evolution of the proposed cellular schemes for mobile radio telephones.

5.6.3 Single vs Four Base Stations

If the number of base stations were limited to a few per service area, then some of the spectral economies of the multi transmitter mode could be realised. This section will explore the number of channels which might realistically (on the artificial infinite hexagonal county plan) be needed to provide one usable channel to a user anywhere in the country (as opposed to the ultimate of 1 for the base station per lamp-post approach).

The starting point will be the hexagonal nesting of service area of Figure 5.6.1. It is reproduced at 5.6.5a with the circle of prohibition shown and a channel number 1 assigned to the centre cell and surrounding cells which just meet the separation criterion. Below it at Figure 5.5.5b is shown the case for four adjacent areas assigned to the same channel (the common channel case with n=4). The prohibited region is now a series of circular arcs as shown, and at first sight the fact that greater area is blighted in this case would indicate poorer spectral useage.

In the following discussion the a. portion of a figure will refer to the single station case and the b. portion to the four station case.

The argument for assigning a second channel for the single station case says:- choose an area adjacent to channel 1 allocation, call this channel 2. Then all other channel 2's are in the same position relative to channel 1's. this is shown in Figure 5.6.6a. The argument for the four station case is similar - choose an adjacent four cell cluster of the same shape as for channel 1 and assign them to channel 2 as in Figure 5.6.6b.

Subsequent channels for the single station case can be assigned in a similar manner to that of channel 2 and the result is that a total of seven channels is needed to cover the country as shown in Figure 5.6.7a. The shape of the repeating seven cell cluster is outlined in heavy marking. When the assignment strategy for the four station case is repeated however it is found that only four channels are needed as shown in Figure 5.6.7b, with the repeating shape again shown in heavy outline.

Now the re-use distance (which has been taken as greater than 2 but less than 3 here) is in terms of the service area, so that if the actual service area of say each transmitter in Figure 5.6.7a were reduced, then the scale of the diagram would be reduced but the relative positions would remain and the same number of channels would be required to cover the country. But, consider now the service areas of the single station situation to be as shown in the diagram 5.6.7a, but reduce the range of those in the four station situation of Figure 5.6.7b by a factor of 2 (ie. shrink Figure 5.6.7b by a linear factor of 2). Now the service area of channel 1 for four stations closely approximates that of the service area of a single station - but only four channels are still needed to cover the country.

Thus, for the situation shown, the number of channels available country wide for a given width of spectrum has increased in the ratio of 7:4 by going to a common channel system.

5.6.4 Spectral Gain of Several Common Channel Base Stations Over the Single One

The techniques shown graphically in the previous section can be extended to cover other values of re-use ratio and, where appropriate, other numbers of base stations to a coverage area.

The geometry of regular hexagons limits the number of hexagons which can group to form repeating shapes without leaving gaps between them - tessellation. Following, the geometry will be described in terms of two numbers i and j as shown in Figure 5.6.8. They are the number of cell centre to centre distances moved along lines at 120 degrees to each other in order to move from one cell centre to another. Thus they have integer values. The cell centre to centre distance - the re-use distance - is given by $D(i^2 + ij + j^2)^{\frac{1}{2}}$ where D is the cell centre to centre distance and the number of cells forming a repeating group is $i^2 + ij + j^2$.

Since the co-channel protection ration Γ is a function of the re-use distance then only discrete values of this are strictly admissible.

Appendix B examines the situation for both the single cell situation and for clusters of various common channel cells. The results are presented in Figure 5.6.9. This shows the relationship of number of channels required to cover the country for the single cell case (and the generating function underlying the stepped graph), and the points calculated for several types of cluster. The nomenclature for the latter is given as a two figure reference n,m. The first is the number of composite cluster structures needed to tile the plane (ie. the number of channels required), and the second is the number of cells forming the cluster.

It will readily be seen that the more complex clusters effect a dramatic reduction in the number of channels required (to allow one channel to be available anywhere in the country) compared to the single cell case to achieve a given protection ratio. Simplifying assumptions were made to devise these latter values which involved neglecting the effect of other than nearest to nearest co-channel cell interaction. The factors not considered were the contribution to the wanted signal by the other common channel transmitters which would tend to increase the wanted signal and hence the protection ratio on the one hand, and on the other hand, the contribution to the unwanted co-channel signals from firstly the other transmitters associated with the cluster of the nearest cell and secondly the surrounding co-channel clusters.

A somewhat closer examination of the 7 cell cluster case showed that there might be a reduction of the true protection ratio by around 4dB. An exact figure could be produced for the geometries of figures B4 and B5 in appendix B but it would be largely of academic interest only.

Appendix C examines the situation for increasing numbers of cells per cluster leading to the case where there is a virtual continuum of common channel transmitters in a cluster.

The appendix shows that as the number of cells becomes large the unwanted signal level outside the cluster becomes arbitrarily small. This appears to be due to the fact that the total radiated power of a cluster decreases with increasing number of cells so that for very large numbers the total power becomes very small and, in consequence, can have very little impact outside the cluster (within the cluster the received power level is maintained).

Thus the intuitive picture of Figure 5.6.4 is confirmed and in practice it should be possible to implement enough transmitters in a cluster to provide any sensible protection ratio for a 3 cluster (3 channel) situation.

5.6.5 Summary

This section has shown that the operation of multiple base stations to convey a signal within an operational area is more spectrally efficient than operation using a single base station. Using this technique only a few channels are required to be available nationally in order to give one operational channel to each coverage area. The minimum practical number is likely to be three.

The spectral economy of multiple base station operation further enhances the case for common channel operation.

5.7 Overall Summary

It has been shown that, of the three techniques studied for providing area coverage, common channel operation from multiple base station sites is spectrally more efficient than either time division operation or frequency division operation of the same sites. It also offers greater operator convenience. The actual form of common channel operation has been left open for the modulation section to discuss and this must take account of the effect of the mobile receiving RF signals from two or more base stations where there will be radio wave interference.

The technique of common channel operation from multiple base stations has been examined in some detail to show that it is a potentially attractive method of achieving the maximum number of available channels in any one area for a given nationwide spectrum allocation; and in this respect it has high relevance for the so called cellular radio concept of mobile radio telephony.

6. Propagation Factors II

6.1 Variations of the Wanted Signal

Section 4 (Propagation Factors I) gave the theoretical basis for predicting the signal levels to be expected at a mobile assuming propagation over a smooth plane earth. This section will deal with the propagation of the wanted signal in practical environments while section 8 (propagation factors III) will cover the impact of anomolous propagation and factors which cause interference to the wanted signal.

Propagation at VHF for the mobile radio situation has been studied extensively by many workers. (e.g. Jakes 1974, Okumura et al. 1968, French 1976, Allsebrook and Parsons 1977)

There is general agreement on the basic properties of propagation, which are the same at VHF and UHF, and the differences lie mainly in the refinements desired to make more and more accurate level predictions for particular circumstances (eg 900 MHz urban operation)(Bajwa and Parsons 1982). The agreed model shows a wide range of expected signal levels for any given range from a base station transmitter. It is necessary therefore to characterise these variations in statistical probability terms.

Over a short distance, some ten wavelengths, the signal can fluctuate widely with variation of 30 dB or more being observed within some half wavelength. This aspect is usually termed fast fading since most workers assume that the mobile will be in motion through this fluctuating pattern and therefore turn a spacial pattern into a time one. Clarke (1968) has shown that a model which assumes that the mobile receives only a large number of randomly scattered signals, and no direct one, will predict that the levels follow a Rayleigh probability density distribution and that measured results fit this very well. A plot of this form is shown, for an arbitary scale in figure 6.1.1. Thus over a short distance the signal is characterised by a mean of some kind (whether mean, median or rms is at the convenience of the measuring system since they are easily related - Appendix D) and the fact that the probability of receiving a signal at some level related to this mean is determined by a Rayleigh probability distribution. For planning purposes it is more relevant to consider the cumulative probability distribution. This is shown in figure 6.1.2, a non linear scale for the cumulative probability is used in order that the relationship should show as a straight line. This is useful when used for plotting experimental results (whether from actual measurements or from a fading simulator) and in order to give equal emphasis to all parts of the curve.

No experimental results are presented since the measuring system discussed in the previous section on propagation deals only with an average figure (as discussed in appendix B) and fading simulators are of particular relevance to both digital data communications and diversity reception which are not dealt with in this work.

In addition to the fast fading there is a slow (ie longer distance) variation of the mean. The same workers have shown that this can be characterised by a log-normal distribution, ie the level expressed in dB has a normal distribution.

Various attempts have been made to produce a mathematical expression which covers the combined effects of the two variations. Parsons and Ibrahim (1983) have shown that the best fit (at least for urban areas) is an equation derived by Susuki (1977) and evaluated by Lorentz (1980). This has the form for the probability density function of the signal strength in dB (Sd) of:-

$$p_z(Sd) = \int \frac{2}{U} \exp\left\{\frac{2}{U}(Sd - S_R) - \exp\left[\frac{2}{U}(Sd - S_R)\right]\right\} \frac{1}{z\sqrt{2\pi}} \exp\left[\frac{-(S_R - S_s + \frac{z^2}{U})^2}{2z^2}\right] dS_R$$

Where $U = 20\ \mathrm{Log}_{10}e$, S_R = rms value of Rayleigh distribution , z = parameter of Susuki distribution, S_s = rms value of Susuki distribution.

Simple theory suggests that propagation losses between a matched dipole at the base station and a matched quarter wave whip aerial on the mobile will be independent of frequency. In practice there is some increase in attenuation due presumably to increased absorption at higher frequencies by foliage etc., and perhaps increased scattering.

The inherent variability of the signal has significant effect on the choice of modulation techniques and of course on the trasmitter powers required.

6.2 Co-channel Effects

The first, obvious, effect of the signal variability is that even when the median level of the wanted signal can be predicted there is a probability that the level found at any one spot could be significantly below this and indeed could be below the minimum acceptable. This is usually set by the receiver noise level and is allied to the setting at which the mute (sometimes called squelch) operates to disconnect the audio output of the receiver - the mute level. Thus no matter how high the transmitter power there will be patches or 'holes' in the coverage pattern where the signal is not usable.

In addition to this effect there will be places where a distant but cochannel signal is higher than would be predicted on an inverse fourth power law - certainly there will be places within the service area where the desired cochannel protection ratio is not obtained due to the wanted signal being below its median and/or the unwanted being above its median.

The effect of the Rayleigh variability on re-use distance was demonstrated by Gosling (1978). He derives a formula for the probability of the wanted signal being higher than the unwanted by a given protection ratio. This is of the form:-

$$\Lambda = \frac{P_1\left(1-\frac{r}{u}\right)^4}{\Gamma P_2\left(\frac{r}{u}\right)^4 + P_1\left(1-\frac{r}{u}\right)^4}$$

where Λ is understood to mean the probability that the field strength of transmitter 1 (of power P_1) at a range r is greater than that of transmitter 2 (of power P_2) by the protection ratio Γ. The distance between transmitters is u.

The nature of this prediction is shown in figure 6.2.1 for transmitters of equal power located at ranges of 0 and 1, two values of are shown. The impact of this will depend not only on the value of (which will be a function of the modulation - to be decided), but also of the nature of the interference due to this effect. Both these aspects will be dealt with in the next chapter on modulation choice.

Suffice to conclude here that even in an ideal smooth earth situation, co-channel effects and their impact are not of a simple nature. This is further developed in the next section.

6.3 The Effect on Re-use

The formula in section 6.2 for probability of receiving the wanted signal can be rearranged to determine the ratio of r to u which will reach a given probability. Assuming equal transmitter powers this is:-

$$\frac{r}{u} = \frac{1}{2}\left[\left(\frac{\Gamma\Lambda}{1-\Lambda}\right)^{0.25} + 1\right]$$

This is of the same form as that to be expected for the simple flat earth situation as given in section 5.6.2, but with the normal Γ term modified. Therefore the concept of a modified protection ratio Γ' can be introduced where:-

$$\Gamma' = \Gamma\left(\frac{\Lambda}{1-\Lambda}\right)$$

Thus the effective protection ratio to be used for calculations where Rayleigh fading is important will be higher than the one assumed or derived from a static situation. The multiplying factor being the term $\left(\frac{\Lambda}{1-\Lambda}\right)$. This can have a marked effect since it is unsual to talk in terms of the probability of achieving communication of 90% or more, frequently 99%. For these the normal protection ratio would have to be multiplied by factors of 9 and 99 respectively.

Before doing this however it would be sensible to make some allowance for other (propagation) factors which will cause the wanted signal to be worse than that desired. These factors could be the shadow fading (log normally distributed) and long distance interference due to anomolous propagation. If the 'probability budget' were split equally between the Rayleigh fading factor on one side and all others on the other then the figures to use for Λ for the two previously quoted examples would be 95% and 99.5% respectively. This leads to enhancements of Γ by factors of 19 and 199 respectively; and is equivalent to adding 12.8 dB and 23 dB to the normally quoted protection ratios!

The effect of this for frequency re-use will be apparent from figure 5.6.9 of section 5.6.4. The number of cells will be considerably increased for single transmitter situations. It further enhances the benefits of the multi transmitter operation as is shown on that same figure.

7. Modulation

7.1 General

Initially the frequency change programme, occasioned by the decisions of WARC 79, was viewed as simply a relocation of operations into a different part of the spectrum. All that would be needed was a frequency shift, and if that were small enough then perhaps the existing equipment could be retuned to the new frequency bands. If this were not possible then some component changes, or 'refurbishment' might be all that was required. This philosophy was modified by the realisation that the Directorate has in any case an ongoing programme of replacing time expired equipment. So that it would be sensible to buy equipment to the new requirements and use these equipments for the initial stages of the changeover, any non-time expired equipment on the old frequency bands which was displaced by this could be redirected to areas currently using even older equipment. There would be at worst only a small economic penalty to pay for this since retuning of the older (non synthesised) equipment or refurbishment would be an inherently costly exercise which could be offset against having to introduce the new equipment before the cost of the old had been completely amortized. [The fairly rapid realisation that the size of the frequency shift would be bigger than the original philosophy would admit, was an additional factor].

Once new equipment provisioning even if only that of the mobile, became the working assumption then any need for compatibility of operation between the pre WARC and post WARC situations was broken. Thus the question of the best modulation to be used could be considered. It will be recalled that modulation is a key element of the interaction diagram of section 3.2. Thus it will have a considerable influence on a number of succeeding elements and will, by virtue of the return loops of the interaction diagram, be fairly directly influenced by a large number of the elements. Furthermore it will be seen from the rest of this chapter that it is a subject which requires considerable internal considerations.

The author, by virtue of his advisory role to the WARC frequency change programme, was asked to advise on the technical aspects of the choice of modulation and provide a recommendation. This fitted well with knowledge gained from various internal studies undertaken within the Directorate and those which had been placed with Universities etc. The knowledge thus gained formed a general background and had never been directly addressed to answering the question posed here.

The chapter is organised in order of both the logical series and the approximately chronological series of events. It discusses the various forms of modulation considered, selects a short list, and, after discussing some basic factors, analyses on a theoretical basis relevant aspects from the stand point of firstly normal single station operation, and then the special factors which come into play for multi-station quasi-synchronous operation. This is followed by a description of a practical evaluation and the conclusions drawn. Subsequently the particular system parameters, which will affect primarily the design of the fixed links, are discussed and subjective tests described which set limits to these design parameters.

Some of the comments and approaches, particularly in the initial sections of the chapter may appear trivial or basic, and therefore some comment is called for. They have been included on two main counts; firstly there is a danger of assuming that the various types of modulation have inherent properties but, in fact, these properties are not inherent they are a consequence of the form or nature of general usage. Secondly some aspects of the mode of operation called for in this discussion will require an understanding on probing of not only the fundamentals but also the properties of the hardware involved. Thus the early sections serve to set the scene against which the later sections will enact the story.

7.2 The Contenders

7.2.1 The Modulating Signal

The message which it is desired to convey from control to mobile, or vice versa, is that of analogue speech. This, then, is the signal which it is desired to impress on the radio frequency carrier as a modulation. It equates closely to a telephone speech signal in that the bandwidth required is from some 300 Hz to 3 kHz+ (although the upper end is not hard and fast). Overall quality is again similar to that of telephony practice - not the individual circuit requirements but the overall end to end performance. This can however be modified; the users of some radio systems, e.g. military networks and airline communications, manage to extract the meaning from signals which are virtually unintelligible to the normal telephone users. They rely on using trained operators, well practised in their job, and probably using a limited set of vocabulary (with possibly a modified frequency response). The emergency services have some commonality with those users but only to a limited degree. Therefore the quality can be allowed to fall below that of the standard set for telephone usage. Having said that, it must be remembered that they are the emergency services and safety of life is frequently at stake. Overall, therefore, the aim is to provide approximately telephone quality with recognised degradations below this to a limited extent. Such degredation can then be considered to be the consequence of the odd spots of poor RF signal reception due to multipath and/or common channel/quasi-synchronous operation.

Quality is a function of a number of factors as will be discussed in later sections of this chapter but the simplest, which also admits of relatively simple design calculation is that of audio signal to noise ratio. A figure often quoted as being acceptble for the emergency services (and other 'professional' private mobile radio users) is 20 dB. This then will be taken as close to the target for performance in the system under consideration. [It is perhaps worth noting that a figure of 12 dB which is often taken as the limit of degradation for certain interference criteria produces, to the author's ears at least, unintelligible communications.]

The question of transmission of digital data signals has been under consideration within the Directorate for some while. For the modestly high speeds (1.2 - 2.4 kbit/sec) under consideration, it has been shown that freedom from having to implement it by imposition onto an audio baseband channel is highly advantageous - in fact virtually essential. Thus it is very unlikely that a mobile radio scheme designed for the emergency services for speech use would be expected to carry this medium speed data in place of the speech. Additionally the data signal would set, probably, very different system design criteria even if it were implemented on the audio channel. For these reasons, plus the fact that the introduction of medium speed data working would be delayed due to the WARC work load, only analogue signals of telephony type have been considered.

The last clause of the last sentence covers the admission of existing digital data signals which are low speed (at a maximum rate of not more than some few hundred bit/sec). These are usually implemented by tone encoders, for supervisory and operational signalling, or via telephony type modems. All these operations can be considered however as being supported by any channel which is suitable for speech.

7.2.2 The Question of Digital Speech

Having dismissed the subject of digital data from our considerations, that of speech in digital form is one to be considered. Moreover if admissable then the previous arguments on digital data would not apply and they could well be deemed an added service probably with little penalty for the systems design. Thus digital speech needs discussion.

It is currently one of the fashionable topics and thus has an aura of glamour and due to its rapid advances elsewhere, it carries an impression of inevitability. This is believed in the author's view however, to be not the case for mobile radio, and perhaps not for radio in general. To justify this it is worth a short examination of its popularity elsewhere. Two main fields are apparent, signal processing, and telephony usage.

The signal processing aspect benefits from the ability to perform complex operations on digital signals by virtue of the power or capacity of microprocessors and very large scale integration (VLSI) devices. Once the signal has been digitised, then the field is open to apply numerous algorithms to produce a range of effects from complex filtering, to specific impairment removal. For best results the digitising rate must usually be high and the processing probably complex. Thus the requirement becomes one of high speed high cost devices, or delay in processing the signal. The areas where the techniques are in use are those which can tolerate these consequences. In general this does not apply to mobile radio, and in any case it does not imply a requirement to transmit digital speech, all the processing can be done if necessary within the mobile, base station, or control.

The telephony aspects stem from the desires to transmit the speech signal in undistorted form over long distances on the one hand, and to implement non mechanical telephone exchanges on the other. Having once digitised the speech signal then it can be reconstituted at intervals in its passage by simple digital devices; this imposes almost negligible distortion if done before the digital signal itself reaches some threshold of corruption. [It is virtually a return to the original communications by telegraphy with true digital repeater stations]. Once in digital form then the signal processing described in the previous paragraph can be used, and one particular form of this is the switching function of a digital exchange.

Both the usages described require high digitising rates (64 kbit/sec PCM for telephony) to achieve sensible quality and can support this rate because their transmission media are not inherently bandwidth limited. [The ordinary telephone line pair has in general been made band limited by the inclusion of loading coils to make a flat (amplitude) response for analogue working; these coils are simply removed for digital operations].

Mobile radio cannot provide the bandwidth necessary to carry such bit rates in a simple form - not even those with the luxury of 25 kHz channels. To carry digital speech in the normal, or in our case the, bandwidth of 12.5 kHz would require either converting the 64 kbit/sec binary signal to a multi level one and/or reducing the basic digitising rate. Multi level signals are 'fragile' in that they suffer greatly from the factors of distortion due to non ideal channels coupled with increased vulnerability to noise. To get a 64 kbit/sec signal into a 12.5 kHz channel (some 10 kHz available bandwidth) would require a symbol rate of some 8 k symbols/sec ie a 3 level symbol.

Turning to the question of lower digitisng rates, then 16 kbit/sec systems are in use (delta modulation with 3 bit companding) mainly by the military and they require 25 kHz radio channels. Such signals are relatively robust but inherently of a quality poorer than can be considered for use here. Other lower bit rate systems are on the verge of leaving the laboratory stage (the $5 vocoder has been 'just around the corner' for some 10 years now) notable here are LPC techniques (linear predictive coding). But again the digital signal is fragile and although it might be possible to fit into 12.5 kHz channels they are both costly and as yet of insufficient performance.

Therefore for the purposes of this study only analogue speech transmission was considered (with the minor exception cited in the next section).

7.2.3 Modulation Forms Considered

The main modulation forms considered were:-

- AM
- FM
- SSB
- Spread Spectrum

and these will be very briefly further commented on here.

AM - amplitude modulation. Simple, full carrier double sideband amplitude modulation was considered here although the Directorate had been associated with work previously for diminished carrier AM and this will be commented on in the section on quasi synchronous operation. Simple amplitude modulation has been used for many years by the Directorate for the mobile sequent of its communications. The demodulator has been a simple envelope detector.

FM - frequency modulation. More correctly this should perhaps be termed angle modulation since not only is it very close to PM - phase modulation - but in fact the form of modulation used is more nearly that of PM. To be precise it is FM with 6 dB per octave pre emphasis at the transmitter. This is used for the usual reason of making the received noise spectrum (after de emphasis) flat and therefore subjectively more tolerable than the otherwise rising noise spectrum for the same total noise power. The use of pre emphasis and the way in which it is implemented with the possibly fraudulent claims of performance are commented on in appendix E (the effects of pre- and de-emphasis on quoted FM performance). FM has been used by the Directorate for the fixed radio links operating at VHF high band to and from base stations, for the personal radio systems operating at UHF, and for certain special users in a car to car mode at mid band VHF. No particular type of demodulator has been called for or is assumed in the rest of this chapter, but near perfect limiting is assumed.

SSB - single sideband. No particular value of carrier is assumed here - anything from full AM carrier to zero was admissable. SSB is often equated with zero carrier SSB, but for mobile radio it seems obvious that some form of carrier be present to provide ALC (automatic level control) and as a frequency reference for the (necessarily) synchronous demodulation. [The RF signals will be subject to doppler shift which would exacerbate any mistuning of transmitter and receiver. Various forms of SSB have been proposed for mobile radio they are, carrier or pilot (diminished) carrier, zero carrier but a pilot tone above the audio band, and zero carrier but tone in (a notch) in the audio band. The form considered here is that of pilot carrier since it was the most available or mature and more importantly could be made most compatible with AM for a change-over situation. (Zaid and Lockhart 1982). The use of an amplitude (or even frequency) companders, as advocated by some, was not considered since it was thought that such techniques could be applied to AM with similar improvements [FM usually has companders and/or limiter (see appendix E) applied].

Spread Spectrum. A less common form of modulation in which the basic modulation (which can in principle be analogue or digital) is deliberately spread in frequency by an imposed spreading signal. At the receiver knowledge of the spreading signal is assumed to de-spread and recover the original modulation. Two main forms of spreading are frequency hopping and direct sequence, the first is self explanatory the latter consists of an additional stage of modulation by phase reverse keying from a spreading digital signal. Such techniques are in use by the military for security and anti-jam reasons and by satellite services for multiple access to non linear transponders. For use by the Directorate they have been studied by Matthews and Drury (1980) and they have shown that a direct sequence system should have about the same spectral efficiency as that of conventional AM or FM if considered on a nationwide basis for the mobile segment. However the system and its equipment is inherently complex, unproven in even experimental form and it was therefore ruled out for further consideration as being not available within the WARC time scales.

Thus AM, FM and SSB are the contenders.

7.3 Basic Factors - Bandwidth

7.3.1 Background

For planning purposes the spectrum available is frequency divided into a number of channels all of the same width - the channel bandwidth. The degree to which this bandwidth, available for each transmission, is practically available depends on a number of factors. When these factors have been considered and quantified then the net bandwidth available can be determined and this can be taken as the bandwidth occupied at any one time by the modulation - the occupied bandwidth. It is of course possible to work in the reverse direction, starting with the occupied bandwidth derive the channel bandwidth, but the latter are always taken as, at least historically, a nice round figure. The channel bandwidth started as 100 kHz for mobile radio and has been successively halved to 50, 25, and now 12.5 kHz. At the moment there is a general desire to talk in multiples of the highest common factor of such bandwidths as used internationally, this is generally taken as 2.5 kHz. Thus 5, 6.25, 10, 12.5, 15, 20, 25, 30 kHz channel bandwidths have been proposed.

Currently the UK is the only administration which has a need to implement channel bandwidth as low as 12.5 kHz, and this is done by effectively splitting former 25 kHz assignments. Such a reduction can have a very significant impact particularly in the case of FM. The Directorate is in the process, in the old frequency bands, of such a change on its AM services and the non Directorate supplied emergency services will have to change from 25 kHz FM to 12.5 kHz on their new modulation. For a combination of largely non technical reasons they indicated a commitment to stay FM.

7.3.2 Channel Bandwith - Occupied Bandwidth

One of the major factors which makes the available bandwidth less than the channel bandwidth is the requirement for practical filters which will allow operation on a wanted channel whilst an adjacent one can be active with much higher carrier level (specifications MPT 1301, 1302). Using the highest rate of cut-off for the normally available crystal filters, this gives a pass band of some 18 kHz for 25 kHz channeling and 7.5 for 12.5 kHz channeling. The band occupied by the modulation must fit into this but not tightly.

If the occupied bandwidth is viewed as a block of width w the situation can be depicted as in figure 7.3.1. The whole of the occupied bandwidth needs to pass through the filter without its edges hitting the sides of the filter response. If it did then AM and SSB systems would experience distortion largely due to attenuation of the baseband frequency components corresponding to the conflicting edge of the occupied bandwidth. FM systems would have distortion largely due to the non linear phase shifts occurring in the region.

The occupied band is uncertain in location due to three main causes of frequency drift:-

- Transmitter carrier frequency drift

- Receiver local oscillator frequency drift

- Receiver filter centre frequency drift

The causes of the drift are mainly ageing/time and temperature.

For 25 kHz system it is usual to assume that each of the above will be some 1 kHz in magnitude and, for 15.2 kHz system each 0.5 kHz. It is too restrictive to consider these as adding on an algebraic basis since they may well not drift in the same sense. So the figures often assumed for the total uncertainty are 1.5 kHz for 25 kHz channels and 0.75 kHz for 12.5 kHz channels. These again are peak figures and mean that the occupied bandwidth can take on the values of 15 kHz and 6 kHz respectively.

For the case of SSB the filter technology and stability of oscillators etc is still under discussion and development.

7.3.3 Required Modulation Bandwidth

It was stated earlier that the message bandwidth - the baseband - was approximately 300 Hz to 3 kHz. Thus for an SSB system which effectively translates the baseband to RF, the occupied bandwidth is at first sight 2.7 kHz. This however ignores the requirement for a pilot or carrier; for this it is reasonable to take the pilot carrier system and say that the occupied bandwidth is ideally 3 kHz. Following on from the last sentence of section 7.3.2 however the available bandwidth from currently practical filters and for 5 kHz channelling admits of only some 2.5 kHz for baseband bandwidth.

For AM systems the modulation will occupy twice the baseband - that is 6 kHz for our goal. This then just fits that available for 12.5 kHz channelling.

The modulation bandwidth of FM systems is undefined. Theoretically it is infinite but most of the energy is within a relatively small range of the carrier centre frequency. So that as the FM modulation is subjected to decreasing filter passbands the demodulated output is increasingly distorted. The distortion is very small and can be neglected for filter bandwidth downs to some given value. This value is of course dependent on the distortion tolerable, and an often used criterion for narrowest bandwidth is that due to Carson and termed the Carson Bandwidth. This is given by

$$w = 2\ (f + \Delta F) = 2f(1 + m)$$

where w is the occupied bandwidth as before
f is the highest baseband frequency
ΔF is the peak frequency deviation
m is the deviation ratio

This fits well with the regulatory parameters for 25 kHz FM which allow a peak deviation of 5 kHz. Thus the expected occupied bandwidth would be 2(3+5) = 16 kHz.

At this point it is instructive to compare the expected performance of a 25 kHz FM system with that of a 12.5 kHz system from the viewpoint of received audio signal to noise ratio for a given receiver. This is proportional to the square of the deviation ratio so that a table can be constructed as in Table 7.3 below

Table 7.3 Comparative FM Performance

Channel Spacing kHz	25	12.5(a)	12.5(b)
Occupied Bandwidth kHz	16	9	?
Peak Deviation F kHz	5	1.5	2.5
Deviation ratio m	1.6	0.5	0.8
Highest baseband frequency kHz	3	3	?
Relative s/n dB	0	-10	-6

This requires some comment. The 25 kHz channelling case is self consistent and provides the reference signal-to-noise ratio for comparison purposes. The column headed 12.5(a) is what would be expected on the basis of the reasoning given above (given a generous allowance for occupied bandwidth!), and it will be seen that the performance is considerably degraded (by 10 dB). In fact the peak deviation allowed by the regulatory authorities is 2.5 kHz as shown in column headed 12.5(b). To equate this with the oft quoted degradation in performance of 6 dB, then the deviation ratio must be 0.8. This yields a top baseband frequency of 3.125 kHz - more than is required. But the occupied bandwidth would become 11.25 kHz - far higher than is allowable. Hence question marks have been put in the table as shown.

To the extent that the discrepancies are ever discussed the lack of precision of the Carson Bandwidth formula is brought into play, also, the relationship of peak speech voltage to its mean, and the relationship of this to any test tone. This is another case of the, at best, imprecision or inconsistency of the FM mode of operation and its proponents here (again reference to other aspects is made in appendix E).

At this stage it suffices to say that both AM and FM can be implemented in 12.5 kHz channelling but users of the latter who have been used to performance at 25 kHz channelling may well notice a significant reduction in performance. The signal to noise ratio will worsen and in all probability the audio bandwidth will have to be reduced.

7.3.4 First Effects of Quasi Synchronous Operation

Previous chapters have commented on the fact that quasi synchronous operation will require very stringent control of the offset of base station transmit frequencies. Later in this chapter it will be shown that this should be less than 10 Hz and preferably, for AM at least, about 1 Hz. Without a feedback connection between the various sites of a county to slave all to a master reference one, then they must be effectively free running and hence extremely stable. This then counters the 0.5 kHz tolerance assumed in section 7.3.2 and makes the transmitter drift allowance effectively zero.

Another factor is that mobile receivers will almost certainly have their channel frequencies determined not by a bank of separate crystals as in the past but by synthesis from a master reference. It is economic therefore to consider that a device with better stability and ageing can be used; therefore the receiver drift allowance can be reduced.

In the practical case under consideration therefore the overall drift allowance can be significantly reduced - a figure of some 0.5 or even 0.25 kHz could be used.

This goes a little way to admitting the figures used for FM in the above table, and it further consolidates the SSB situation. For the AM case however it means that a little more baseband bandwidth could be tolerated, say 3.5 kHz. The extra bandwidth could be used to enhance speech clarity or perhaps more benefically to carry the low speed data discussed, at the start of the chapter, as an overlay.

7.3.5 Summary

This section has shown that operation in 12.5 kHz channelling situation is viable for AM and FM. The quasi-synchronous mode of operation imposes constraints on equipment which are beneficial in ultimately improving baseband performance. But whilst FM users must expect a significant degradation compared to their use of 25 kHz channels, AM users may in fact gain some benefits such as a low speed data overlay.

The situation for SSB operation in 5 kHz channels has not been discussed in detail but reliance has been placed on its advocates being able to achieve acceptable performance. To this end both amplitude and frequency companding have been proposed.

7.4 Single Base Station Case

7.4.1 General

This section begins the comparisons between AM, FM, and SSB. The first two at least have been effectively opponents in a modulation battle since the inception of FM. SSB has entered the fray on certain sites and from time to time. The main battle ground has been that of broadcasting with the general outcome that AM was favoured wherever device simplicity or bandwidth constraints applied. FM gained if fidelity (particularly linearity) was required and/or bandwidth was available to be traded against transmitter power or receiver noise. The arguments are well known and are not repeated here.

However there are particular facets which bear on mobile radio operation and which are perhaps peculiar to it. The question of multiple base station operation will be dealt with in the next section but even the case of what, to the majority of private mobile radio operators is the norm, single base station operation, has some pertinent aspects which need to be taken into account in the comparisons. After all even in the multi base station situation, much of the coverage of a county can be considered as having a dominant base station.

The comparisons will therefore be considered under several headings and the reasonings used will be more descriptive and comparative than quantitative (again reference can be made to the extensive literature and standard text books for the details).

7.4.2 Gross Range and Slow Fading

The variation of signal level due to the distance between a mobile and the base station has been discussed for the smooth earth case in chapter 4 and this extended the argument to firstly the case of slow fading. The main import of this was the large dynamic range of RF signal which the receivers (both mobile and base station) would have to encompass.

For FM, figure 7.4.1 shows the theoretical output (audio) signal to noise ratio (s/n) as a function of RF input carrier level (C) with deviation ratio, or its consequence - channel bandwith B as a parameter. The RF carrier axis is in fact shown in terms of the carrier-power to noise-power-density ratio (C/No). This normalises the scale making it independent of the receiver bandwidth, thus enabling true comparisons to be made between systems of different bandwidth. No actual values for the axes are shown since this is not important at this stage, only the comparisons. Additionally there can be considerable debate on what constitutes the audio signal power (s) and how and under what conditions the audio noise power (n) is measured. The figure shows the conventional feature of, for a given channel bandwidth, the output signal to noise ratio being proportioned to input carrier level down to a certain carrier level. At this threshold point (approximately 10 dB C/N where N is measured in the IF bandwidth) the audio s/n drops very rapidly as the carrier level is further reduced. If the deviation of the carrier is increased then for a given carrier power input the output s/n in increased - providing operation in above the threshold. If below it, the converse is true, and the threshold is worsened as the occupied bandwidth is increased.

Figure 7.4.2 shows the same situation with the curves for AM - (assuming an envelope detector) and SSB added - again high precision is not aimed at for the same reasons coupled with those of the differences between theoretical and practically used modulation depths/indexes and whether carrier powers are equated on a peak or mean basis. Also shown on the figure is the approximate position of 20 dB s/n - the desired target.

From this figure it will be seen that AM and FM in 12.5 KHz channels give very similar performance - above the latters threshold, with SSB as to be expected some 3 dB worse. If the 20 dB figure is taken as a minimum objective then the corresponding carrier power can be deduced and used to determine the transmitter power for this figure to be achieved at the edge of service. In practice a margin is added to this to cover the shadowing/slow fading and some of the fast fading - say 20 dB.

A point of interest arises from this figure. If the minimum acceptable output signal to noise ratio is as shown then, when the carrier level is higher the signal is better, but this enhancement may not be of real benefit. If 20 dB is adequate then anything more is perhaps a waste.

There are two implications from this. The first is that area coverage by multiple base stations is again seen to be beneficial since the range of carrier powers over the county is reduced to that which is just needed. The second point is that the improved signal to noise ratios obtained by operating with FM in bandwidths greater than 12.5 kHz is also difficult to justify - 12.5 kHz is adequate and due to the threshold effect no extra range is conveyed by the extra deviation!

7.4.3 The Need for Limiting, ALC, and Mutes

The FM audio signal level is maintained at a constant figure, despite variations in input RF carrier level caused by fading etc to the limiting action of the receiver. This is usually viewed as 'clamping' the level of the carrier at intermediate frequency before (or even as part of) the demodulation. What happens in practice is that the carrier is subject to a constant high degree of amplification before reaching the limiting stage at which point it is often clipped to a set level. This clipped level is then possibly filtered and used as the input to the demodulator. The clipping, acting on the IF carrier, is effectively instantaneous as far as any carrier level variations are concerned so these do not reach the output. (It should be noted that there are phase variations due to fading, since they cannot be removed their effect will be present in the audio output.)

The clipping/limiting action is maintained for carrier levels down to about the threshold region. Below this level any input level variation would appear as output signal level variations.

Since the output signal for carrier below the threshold is, for all deviations, considered virtually unusable, then a mute circuit is brought into operation. This senses the carrier level and disconnects the audio output if the level is too low. The need for a mute arises from the fact that the received signal might be too low due either to poor carrier input (range, fades etc), or to the fact that the base station is not transmitting, Without the mute the listener would be subject to a high audio noise level under these circumstances. This is different to the broadcast situation where any signal fades are of relatively low depth and the transmitter only switches off at the end of the day.

The AM and SSB receivers need similar facilities. Again the fading characteristics impose amplitude variations on the carrier which would appear to be in possible conflict with the modulation. The answer lies in the fact that the rate of variation due to the causes so far described, and those due to the fast fading and quasi synchronous effect, are much slower than the rate of variations due to the lowest modulating frequencies. Thus it is possible to sense the AM carrier/SSB pilot level and use this to control the gain of the amplifying stages such that the level is maintained constant, hence automatic level control (ALC). Unlike the limiting action of FM this then is not instantaneous but has associated with it a response time determined by the lowest audio frequency it is desired to pass. The demands on the ALC control loop are severe in that a large dynamic range is desired and as fast a response as possible. Sometimes a two stage loop or two time constant one are employed. [It is worth noting in passing that the rate of fast fading is proportional to the carrier frequency. This means that AM systems and SSB pilot carrier ones are probably limited to operation in the VHF region - operation at UHF at 450, 900 MHz and above would be very difficult].

Muting is employed in AM and SSB receivers for the same reasons as for FM ones. The circuits for corrections of carrier level, and the muting action, work reasonably well in current systems. [Although McGeehan (1981) have shown that poorly designed ALC systems can in fact be deleterious]. But they have two areas where their effects are undesirable. One of these is where the mobile is travelling through a fading pattern which brings it to below the ALC and mute thresholds, the effect is similar to operation in a two (weak) signal overlap area of a quasi synchronous system and will be discussed in section 7.4. The other situation is that occasioned by the end of message point when the mobile or base station transmitter, switches off. The details of the effects here are given in the next section.

7.4.4 The Effects of Limiting, ALC and Mutes at End of Message

This section will describe the action of ALC and limiting on an unmodulated carrier and compare the outcome for the contending modulations.

Consider firstly the AM case of a carrier of amplitude R volts at the input to the receiver where the noise density (due to all causes) is N_0 watts/Hz. The noise bandwidth of the receiver will effectively be determined by the IF stages and filtering, and this can be taken as w Hz (close to the passband or available bandwidth).

The carrier to noise ratio at IF will then be

$$\frac{JR^2}{wN_0}$$

where J is a proportionality constant taking account of the units of measurement, impedances etc. This carrier to noise ratio is maintained irrespective of any attenuation or amplification within the receiver. The amplification will in fact be controlled by the ALC circuit which attempts to maintain the values of R constant at the output of the IF stages - the input to the detector. This is achieved by making the gain G of the preceding stages inversely proportional to R.

If the amplifier gain is considered to operate on the carrier component and the noise component simultaneously but separately, (this will be the situation where the amplifier is linear and the carrier detection circuits of the ALC are not corrupted by noise) then the carrier to noise ratio will be

$$\frac{V_c}{V_n} = \frac{G^2R^2}{G^2wN_0}$$

(the proportionality constants will be dropped from here on since they serve only to equate units of measure and are not fundamental). Where Vc and Vn are carrier voltage and noise voltage respectively.

Since G $\sim \frac{1}{R}$ or ignoring constants $= \frac{1}{R}$

then :- $V_c = \left(\frac{1}{R}\right)R = 1$

and :- $V_n = \left(\frac{1}{R}\right)(wN_0)^{\frac{1}{2}}$

and the noise power $P_n = V_n^2 = \frac{wN_0}{R^2}$

Thus the effect of ALC is to make the IF noise power inversely proportional to the carrier power (R^2). Assuming that the demodulation process is linear, then the audio signal will be held constant and the audio noise power will be proportioned to the IF noise power. Thus any variation of input carrier signal in AM and SSB receivers will appear as baseband noise power variation in the inverse sense.

For the FM case the outcome is the same although the mechanism of achieving it is different. The baseband signal is inherently constant since it is determined by the deviation of the carrier and is therefore unaffected by the attendant noise. But smaller carrier levels allow the noise deviation to form a larger portion of the total deviation so that the baseband noise power is inversely proportioned to carrier power level (Fagot and Magne 1961 etc).

Thus, so far, the linear and quasi linear modulations (SSB and AM) behave in the same manner as the non linear one FM. So that, for any of the modulations, as a mobile, say, traverses an unmodulated fading field pattern the effect will be heard as a variation in the audio noise level. With the peaks of noise ocurring at minimum of the field and vice-versa. This is shown in the left-band portion of figures 7.4.3 and 7.4.4 for AM/SSB and FM respectively, where the upper line represents the carrier and the lower shaded portion is the audio noise.

The difference between the modulations occurs when the transmitter is switched off as indicated by the fall of the carrier level to zero.

For the AM/SSB case the ALC circuit senses the absence of carrier as a 'low level' and increases the gains accordingly, but there is the inherent time constant associated with the control loop so that the response is not instantaneous (it is however fast enough to cope with virtually the fastest rate of fading as discussed previously). The amplifiers will not in general reach full gain before the mute circuit (after its built in delay) senses the absence of carrier and terns off the audio. Thus there will be a noise burst of increasing magnitude at the end of a transmission. This is termed the 'mute tail', and is shown on the right of Figure 7.4.3.

For the FM systems it can be considered that the limiter is in operative, so that the full RF and IF gain is given to the noise alone. Thus its level, and in consequency the baseband level, is at a maximum. But in contrast to the AM/SSB case the higher gain and level is achieved instantaneously. So that the mute tail is at a constant high level until the mute (this again has a time constant associated) disables the audio stages.

The differences between AM/SSB operation and FM operation can then be seen from the right hand side of figures 7.4.3 and 7.4.4. The difference are particularly noticable when the mobile has been receiving a strong carrier. For the AM/SSB case the mute tail starts at a low level and is cut off before it can reach a very high one, whereas the FM receiver changes from a good signal to noise ratio suddenly to a very high level of noise - perceived as a raspberry! This is in itself disturbing, but tends to be ignored by FM mobile radio users due to acclimatisation - at least for those with 25 kHz channelling or greater. When 12.5 kHz channelling is used the effect will be significantly greater. This is perhaps best explained by considering first operation with a 25 kHz receiver where the receiver volume control will be set for a suitable audio level. Now reduce the deviation of the transmitter to that of a 12.5 kHz scheme, and the volume control setting would have to be increased to give the same audio signal output. The inherent mute tail in each case will be the same (and very little different when 12.5 kHz filters are fitted in the second case) so that after the increase in audio gain it will appear correspondingly louder - 6 dB so if the right hand column of Table 7.3 (section 7.3.3) is considered.

Thus in the emergency services mode of operation where the mobile inherently turns off its carrier at the end of transmission and the base station choses to do so for other reasons, the listener will be subject to a burst of audio noise at the end of each message. This is significantly more disturbing for 12.5 kHz FM systems than for AM or SSB modulation. For a user on a tour of duty of some 8 hours this effect alone could be very tiring.

7.4.5 Co-channel Effects - Capture

The phenomenon of capture in FM system is more an item of the associated folk lore and general jargon than of quantitative definition. There are definitions for the wide deviation broadcast case, but not for mobile radio. In the former case figures of 1dB or less are quoted.

The nature of the capture effect is that, when a FM receiver receives two (differently) modulated carriers of different levels, then the modulation on the stronger carrier dominates the output audio signal of that receiver. If a criterion is established to measure the degradation of the stronger carriers modulation (say a SINAD measurement) then the ratio between the stronger and weaker carriers to first achieve this degradation is called the capture ratio. The picture painted by the FM lobby to the lay user is usually to imply that the capture is complete and instantaneous. That is, only the stronger is noticed. Even if this were true then the effects at the edges of coverage due to a co-channel transmitter described in section 6.2 are ignored.

Experiments were made by the author and a collegue some time ago to measure the capture ratio for mobile radio use and a definition was generated. This gave figures of some 6-8 dB for a SINAD degradation from 20 dB to 12 dB. Even with a carrier ratio of 20 dB the influence of the unwated carrier could be distinctly detected. The capture ratios quoted above are for 12.5 kHz channelling system - the capture phenomenon is dependant on the wanted carrier deviation so 12.5 kHz figures are significantly worse that the old standard of 25 kHz. One of the interesting results of this work was the realisation that when the carriers are separated in amplitude by less than the capture ratio, the output signal is unintelligible.

The effect of this can now be discussed for the field of service range/frequency reuse.

Figure 7.4.5 is a composite one and follows those of section 5.6 showing received signal level as a function of range from two base stations radiating different modulations on the same frequency. Also shown are the protection ratio for AM and the capture ratio for FM Thus the service ranges for AM is shown in figure 7.4.5.b; the idealised ranges for FM assuming perfect capture at figure 7.4.5.c, and the FM ranges assuming a more sensible figure for capture at figure 7.4.5.d. This appears to show that FM would have improved range or improved reuse compared to AM.

Two important effects come into play here, however, in combination. These are the blurring of the concept of a sharp range figure by virtue of the Rayleigh fading as described in section 6.2, and the fact that, in the FM case, in the gap between ranges so far discussed, the signals are unusable; whilst in the AM case there is superposition of the modulations - relatively undistorted. The effect of these two is that the determination of range now should be on an acceptability criterion and this takes the form shown in figure 7.4.5 e and f. Again quantative figures are not attempted due to the highly subjective nature of the measurements, but it will be seen that the ranges are almost equal and there is the choice of better performance for FM close to the transmitter to be traded against a more rapid fall-off at extreme range.

7.4.6 Co-channel Operation - Punch Through

It is usual when considering communication systems which are subject to interference, to divide such interference into two main categories; intelligible, and unintelligable. The latter is usually roughly equated to noise (random, white, gaussian), and the tolerance of a user to it assumed to be the same as that to noise. Intelligible interference on the other hand is considered to be far more disturbing for the same level of noise - thus the level of intelligible interference is set lower than that of unintelligible. This is defendable in the case of telephone operation when intelligable interference is disturbing by virtue of drawing the listners attention away from his conversation, and additionally has privacy implications. This sort of reasoning is usually carried over to the mobile radio scene and used to the benefit of FM. This subsection will show that the findings in fact reverse that conclusion - certainly for the mode of operation of the emergency services.

Consider the case of a mobile, at a moderate distance from a base station, transmitting a message. This is represented diagrammatically in figure 7.4.6, where B is the base station and M_1 the mobile. Another mobile at M_2 starts transmitting (this should be inhibited by him hearing 'pip-tone' but there are several reasons why this might be ignored - as will be discussed later). Now, if AM or SSB is in use, then the controller will hear the second mobile's call sign, admittedly at a lower level, and he can by virtue of the 'cocktail party effect' choose to listen to either of the transmissions and perhaps take some account of both since they will be undistorted. If the mobiles are radiating FM however then one of three situations obtains depending on whether the received carrier strengths differ by more than the capture ratio or not:-

a) M_1 much greater signal than M_2 - M_2 not heard

b) M_2 much stronger tham M_1 - M_1 not heard

c) M_2 within the capture range of signal from M_1

- result neither signal heard, corruption.

Now the reasons for M_2 transmitting are many but the two most important ones are:-

i) accidental transmissions, due probably to a fault

ii) desire to pass a very urgent message.

Both are critical. The former since it can be a long term disuption and the latter by the nature of the prompting incident.

For FM then there is a range annulus around the base station which contains one of the mobiles and extends to ranges equivalent to the capture ratio either side of it. This is shown in figure 7.4.7. Within this range neither annulus transmission is heard, outside it M_2 is not heard and inside it M_1 is not heard. This is a situation which ranges from unacceptable to disasterous since both messages, or their import, are needed.

For AM/SSB there is no such situation the second mobile should be heard - albeit at perhaps very low level if at extreme range.

This factor is important for the emergency services!

7.4.7 Summary of Single Base Station Case

The operation of a mobile radio single base station scheme has been examined in some depth, in the areas where there could be differences, to compare the performance for the three contending modulations AM, FM and SSB.

It was found that, as was to be expected, AM and SSB could be largely coupled together for similar performance and that the differences in general between these two and FM were not significant, being qualitative rather than quantitative - at least for 12.5 kHz channelling.

There were two factors however where there was a significance difference between the first two 'linear' modulations and the non linear FM case. The first of these was the phenomenon of the mute tail, which although at first sight could be considered trivial, would become a major nuisance if an operator were subjected to prolonged exposure to it. The second was the phenomenon termed 'punch through' *which is of particular interest to the emergency services, although the occurrance of the situation leading to this phenomenon is not frequent, the effect is such as to give higher marks to the AM/SSB mode rather than FM.

* 'Punch through' can be defined as the ability of a message on a weaker carrier to be understood when listening to another message on a stronger carrier. It is similar to, and in this case relies on, the 'cocktail party effect'.

7.5 Common Channel (Quasi-Synchronous) Case - Theoretical

7.5.1 General

The contents of this section, and some succeeding ones in chapter 7, have much in common with a publication of the author (Fudge 1984a) which was derived from this study. However the work is reproduced here. in somewhat different form, for several reasons:-

a) At the time of writing this, the publication date of the paper quoted above had not been determined.

b) The publication deals only with the AM case.

c) This section contains extensions not dealt with in the publication - in addition to FM and SSB.

d) It enables the topic to be dealt with as a whole in this text.

The system investigated is shown diagrammatically in Figure 7.5.1. A mobile is in the overlap area of two of the base stations (Bp and Bq) of the system. It will be shown later that this is more critical than for an overlap area of more than two transmitters. The base stations radiate nominally the same signal (modulation) which is derived from a common source at control but reaches the base stations by traversing paths which are different in terms of length, number of stages, nature of carrier (eg line or radio relay), equipment types and their alignments. This results in a modulation mismatch which can be characterised in terms of amplitude and phase or delay. The carriers received from Bp and Bq will have slightly different frequencies and can have a range of amplitude distribution dependent in practice on the position of the mobile both in gross terms of range, and in fine detail where the effects of shadowing and multipath are apparent.

7.5.2 Unmodulated Carriers - AM and SSB

The amplitude of the carrier from base station Bp will be designated P and this will be taken as having the reference phase angle (ie zero), and the amplitude of that from base station Bq will be Q with a relative phase angle of α. Figure 7.5.2 represents a phasor diagram of the situation with the two signals combining to produce a resultant $R = (P^2 + Q^2 + 2PQ \cos\alpha)^{\frac{1}{2}}$

Due to the difference in frequency between the carriers the angle α will be varying continously, with the consequence that the resultant amplitude (envelope) varies with time in the form shown in Figure 7.5.3 which shows the two cases where Q is 1dB (solid line) and 3dB (dashed line) below P; this waveform will appear at the output of the demodulator.

It will be seen that it is rich in harmonics, and increasingly so as the two RF signals come close in amplitude. When they are in fact equal then of course the resultant falls to zero at one instant. The harmonic content of the waveform has been analysed (Al-Nuami 1978 and 1981). There have been proposals to place a notch filter in the mobile receiver, tuned to the carrier offset frequency, to eliminate this beat note, but the diagram and the analysis show that significant amounts of harmonics would remain. There is also the possibility that the filter may not remove the fundamental subjectively, due to the phenomenon used to effect in constructing very low frequency organ pipes (Wood) where the illusion of the presence of a fundamental is created by sounding only its harmonics.

For both the AM and SSB cases the ALC circuits will act on the instantaneous carrier, and, if the rate of variation is within its range, will ensure that the variations are eliminated before reaching the detector. So that in fact the analysis by Al Nuami does not conform to the real situation. Instead it is the noise component which is made to vary in inverse sense - as described in section 7.4.4. Thus the demodulated audio output appears as in Figure 7.5.4.

It should be noted that, whilst the waveform shown in Figure 7.5.3 would in the absence of ALC actually appear as an instantaneous voltage at the demodulator output, the waveform shown in Figure 7.5.4 represents the power envelope of the wideband noise which appears in the audio stages. If the offset frequency between carriers is low, then this is perceived as a rythmic variation in noise background - a breathing effect - at the rate of the offset frequency and with a depth depending on the ratio of carrier amplitudes. At high offset frequencies it becomes a buzz. In either case the inclusion of a notch filter tuned to the beat frequency which has been advocated, or even a high pass filter (which is inherent in mobile receivers) will have only a limited value since the noise spectral components extend over the whole audio bandwidth.

7.5.3 Unmodulated Carriers - FM

As far as the carriers are concerned the situation is as depicted in the phasor diagram of Figure 7.5.2. The resultant amplitude of the carrier will be the same as for the AM/SSB case and, by the mechanism described in section 7.4.3 for FM, the output audio signal will have the same form as shown in Figure 7.5.4. It will be recalled that this is basically due to the variation in carrier to noise ratio at the RF input.

For angle modulation systems however there is an additional audio output. This derives from the inherent phase/frequency variation of the resultant with respect to either of the carriers alone or to the mean of its phase/frequency. The output of a phase demodulator will be dependant on the angle θ shown in Figure 7.5.2, and can be considered proportional to it. Appendix F derives the equation governing the values of θ and the result is depicted in figure 7.5.5 where the solid and dotted lines represent carriers differing by 3 and 1dB respectively. The actual magnitude of the waveform compared to a reference modulating signal will of course depend on that signal's phase deviation. It will be seen that the peak phase angle will not exceed 90° - provided the larger carrier is taken as a reference. But as carrier equality is approached the waveform becomes steeper and hence of higher harmonic content.

For a true FM receiver the analysis of appendix F shows that the output waveform will be of the form shown in Figure 7.5.6 where the solid and dotted lines have the same meaning as before. The vertical axis of the figure is shown in terms of frequency magnification - the value from this is used to multiply the difference frequency between two carriers to obtain the effective value of the output voltage.

Take for example the carriers which would normally be modulated with a tone giving a 2.5 kHz peak deviation, then, if these (now unmodulated) carriers were 100 Hz apart in frequency and 1 dB in amplitude, the output waveform would reach a peak of nearly 9 x 100 = 900 Hz. This is a peak signal to peak unwanted value of some 9 dB! Again the waveform is rich in harmonics.

These waveforms will also be present in the audio output of an angle modulation receiver. If the receiver is an FM one with de-emphasis, then it is usually considered to equate to a PM one; but the de-emphasis characteristic is usually achieved by integrating the audio signal - or placing a continous amplitude slope filter of - 6 dB per octave in the audio stages. The response of such a filter is unlikely to be maintained below some few hundred Hz. So that the waveform of Figure 7.5.6 is close to that predicted - with some attenuation of the higher harmonics by virtue of the filter.

7.5.4 Modulated Carriers - Identical Modulation

Consider first the AM case where the modulation envelope emanating from the two base stations is a sinusoidal one, with the modulation frequency, depth, and initial phase angle identical. The mobile is roughly equidistant from the base stations so that there is no significant difference in envelope delay but the carrier amplitudes and phases are not identical. The phasor diagram is shown in Figure 7.5.7 where m is the depth of modulation. If the carrier, upper sideband and lower sideband are analysed separately then the carrier component is a duplicate of Figure 7.5.2 and each sideband is again a duplicate of Figure 7.5.2 with a scale factor of $\frac{m}{2}$ and a phase angle relative to the carrier of ϕ. Thus $\phi_1 = \phi_2 = \phi$ and $k_1 = k_2 = m$. The resultant is therefore modulated identically to the individual transmissions and there is no modulation distortion, the noise beat remains however.

For SSB the diagram of Figure 7.5.7 can be used with one sideband removed and the carrier amplitude can be set to equal that of the pilot. It will be seen that the resultant modulation is just a scaled version of either of the originals and that there will therefore be no distortion.

For FM the diagram of Figure 7.5.7 could be modified to show each sideband of the modulating tone but that is not necessary. By the argument of the first paragraph of this sub-section each sideband will be a scaled replica of the original and all phase shifts will be the same. So once again the result will be a scaled version of the original and no distortion will result.

Thus it can be shown simply that, under perfect conditions of modulation match from the base station, all the contending modulation will perform as if only one base station were being received - ignoring any of the effects discussed in the previous sub-section.

The next sub-section will consider the cases of mismatched modulations. The modulations are again considered to be tones and the tones will be of the same frequency. The mismatch may be either: inequalities of amplitude, or phase or of course both. The situation could be analysed rigorously for the whole range of inequalities but it will be shown later that this again is unproductive when other effects are considered, so one case of each form will be shown to illustrate the nature and magnitude of the effect.

7.5.5 Modulated Carriers - Modulation Amplitude Mismatched

Taking the AM case first, then both transmitters are modulated with a single frequency tone but transmitter B_q has only half the depth of modulation (6 dB down) compared to transmitter B_p. The phasor representation is shown in Figure 7.5.8 for the instant where the Q phase is at 90^o to the reference and the carriers are nearly equal in magnitude. It will be seen that the lower sideband of the resultant has been advanced in phase, and the upper retarded by the same amount.

The modulation is therefore no longer pure AM but has phase/frequency components. This angle modulation will be ignored by the AM detector but the demodulated output will no longer be a pure sinewave. However the distortion is not high, particularly in view of the gross mismatch conditions.

The SSB situation can again be considered as a sub-set of AM with similar conclusions in that there will be some distortion but not of high significance.

For FM, the diagram/figure 7.5.9 applies and only two sidebands (of possibly many for single frequency modulation) are shown for clarity.

It could be considered that this then represents narrow band FM. The situation is again examined for the instant when the carriers are in quadrature and the result is similar to that for AM in that the sum of the resultant sidebands no longer describes a locus at 90° to the carrier. Thus the pure FM will be translated into part FM part AM. The amplitude modulation will be removed by the limiter (it can be assured that the AM/PM conversion of this device is negligable for single audio signals) leaving a slightly distorted FM waveform.

In conclusion it can therefore be stated that even under fairly gross amplitude mismatch conditions the distortions suffered by each of the contending modulations will be noticeable but far from disastrous. So that some attention will have to be paid to this parameter but even allowing only a moderate contribution from this cause to the overall distortion budget is not likely to prove onerous in practice.

7.5.6 Modulated Carriers - Modulation Phase Mismatch

Amplitude modulation will be considered first and Figure 7.5.10 represents the situation. This is drawn for the case where the modulation phases differ by 90° and the carriers are depicted at the place or time where they are in quadrature - the modulation depths are identical. It will be seen that the resultant carrier has the same amplitude as in that studied in the last sub-section. ($\sqrt{2}$ times the individual ones) but the main effect is that one of the side bands (the lower in this case) has been cancelled. The modulation is therefore SSB, but it is not SSB throughout the beat cycle (the carrier difference frequency). When the two carriers are in phase then it can readily be seen that both resultant sidebands will be present - although of slightly diminished amplitude. At this point in the cycle the modulation depth will have been reduced compared to the originating ones. The situation for the place/time of half a cycle later than that of Figure 7.5.10 is shown in Figure 7.5.11. This is again SSB but this time it is the upper sideband which is cancelled leaving the lower one.

Thus there will be noticeable distortion to AM reception in this case, although the modulation is in fact upheld despite one or other of the sidebands cancelling. This effect is duplicating that which occurs by design in some SSB transmitters, although here two separate transmitters are involved and the signals summed after radiation. It is a phenomena exploited by Gosling (1976) in a technique called sideband diversity where the modulation phase shift is deliberate and two single sideband receivers are employed with the best output selected - since only one will be in a fade, caused by quasi-synch operation, at a time.

Our case of normal SSB operation can be readily derived from that of the AM situation just described with, of course, one of the sidebands removed. In this case the sideband applitude will be enhanced at one of the quadrature carrier conditions, and the amplitude zero at the other quadrature situation.

For FM the phasor diagram is shown at Figure 7.5.12 for the narrow band version where the modulations are 90° out of phase at the place/time where the carriers are in quadrature. The result, not unexpectedly, is single sideband, and again half beat cycle later only the other sideband will be prsent. It can reasonably be assumed that an FM receiver receiving SSB will have a more distorted output then an AM one receiving SSB. Furthermore the second order sidebands of the FM signal (which are not shown on the diagram) will have in effect double the modulation phase shift and so they will reinforce and cancel at different parts of the beat cycle to the cancellation points of the first order sidebands. Thus the situation will be complex in a practical frequency modulation situation.

Predictions have again been made (Al Nuami 1978 and 1981) for the harmonic content of the ouput of AM and FM receivers under these conditions but the arguments of the next sub-section will show that such findings are of little use in determining practical performance. That sub-section will address the variation of effective modulation depth during the beat cycle for each of the contending modulations.

7.5.7 Modulated Carriers - Near Equal Carriers

The previous sub-section showed that the effect of modulation mismatch - at least in phase - resulted in a variation in the effective depth of modulation during the carrier offset frequency beat cycle. This was demonstrated explicitly only for 90° difference in modulation and mainly for the cases of carriers in phase or in quadrature. This sub-section will show that the modulation depth variation phenomenon is not limited to this degree of mismatch and is more severe for the remaining part of the cycle - that of carriers in antiphase.

To demonstrate this consider the case of an AM system with modulations slightly mismatched in phase and the places/times when the carriers are either in phase or in anti-phase as depicted in Figure 7.5.13 - the carriers are nearly but not quite equal. For the in phase situation, it can be seen that the resulting modulation depth is slightly less then that of either of the individual modulated carriers. When the carrier are in anti phase however the result is very different. The carriers nearly cancel but the sidebands do not cancel to anywhere near the same degree, in addition they have suffered a near 90° phase shift. Thus at this point the modulation depth will have increased considerably and as depicted in the diagram the modulation will exceed 100%. Thus there will be considerable distortion for an envelope detector and this will be of a high level. Whether the additional modulation phase shift would be noticed is a moot point.

SSB, which necessarily uses a synchronous detector, will not suffer these distortions but the variation in depth of modulation through the cycle will be noticeable, the output could be zero for the particular conditions of the last sub-section. There will therefore, with SSB, be a 'cadence' given to the output audio signal at the beat frequency. This is almost certainly more tolerable than the distortion to be expected on AM.

The equivalent narrow band FM situation is shown in Figure 7.5.14. The modulation phase shift will again be seen, and also the greatly increased phase deviation, not to mention the high degree of amplitude variation of the total resultant. The degree of distortion does not at first sight appear as large as that which might be expected in the equivalent AM case, but attention is drawn to the waveform of first the output of a phase demodulator and secondly an FM one. The critical area is where the resultant sidebands are at 90° to the position shown in the diagram (ie a quarter of a cycle of the modulation before or after that shown in the diagram). The sideband resultant will sweep very rapidly across the carrier tip - more rapidly than for true sinusoidal PM. The equivalent frequency deviation will in consequence be high. Thus the resulting waveform will be very non sinusoidal and rich in harmonics. This area is also critical from another viewpoint, that is if the two carriers should come closer in amplitude or even swop dominance. If this were to happen for FM then there could be a sharp spike of output as the total resultant passes to the left of the origin shown in Figure 7.5.14 rather than the right - although the spike will be large if the total resultant is only just to the right also. [This is the phenomenon extensively studied for FM operation below threshold, the field of threshold extension demodulation, and in particular phase locked loop operation].

In addition to this distortion there will be, for PM/FM, variation of the noise during the modulation cycle, this arises from the fact that the total resultant carrier is no longer constant so that the carrier to noise ratio will also vary. This does not happen for AM/SSB since the ALC operates on the carrier component only.

In summary therefore it can be said that mismatches in the modulation of two transmitters will be most noticeable when the two carriers are of near equal amplitude (irrespective of absolute level). All the modulation forms considered will suffer distortion. That imposed on SSB may well be more tolerable then AM or FM and the latter will probably be more vulnerable to such distortion then the former. These statements are made somewhat tongue-in-cheek since they have not been backed by a rigorous analysis but the reasons for this will become apparent later.

7.5.8 Low Carrier Levels - ALC Effects

The modulation distortion described in the last section was independent of the absolute level of the two original carriers - only the difference between them was considered. When the carrier levels are low however, then there will be additional factors coming into play. These arise from the same mechanism as that described in sub-sections 7.4.3 and 7.4.4 and are a function of ALC, limiting, and the mute action. The action of ALC on an AM quasi synchronous signal will be described first and the case of carriers only considered for clarity. The resultant carrier amplitude will have the form shown in Figure 7.5.15 (which duplicates Figure 7.5.3) but also shown on the figure is the level below which the ALC system ceases to operate. For levels above this limit the audio output will take the form shown in Figure 7.5.4 but below it the noise level will be constant whilst the signal level will change. This is shown in Figure 7.5.16 which has been dervied for the case of two carriers 1 dB different in amplitude. The ALC is assumed to give a relative signal output (from a modulated carrier) of 30 dB, and the noise variation corresponds to that which would be expected from a situation where either carrier alone would yield an output signal to noise ratio of 22 dB as shown by the dashed lines which represent single carrier operation.

[It will be seen from the figure that for 2/3 of the beat cycle the overall signal to noise ratio is better than for one carrier alone. This is easily seen from the phasor diagram of Figure 7.5.2, and has been used by at least one author (Murasko 1978) as demonstrating the advantage of QS operation (in that case QS AM) since it is better for more than half the time. It does however ignore the effects of the remaining one third of the cycle].

The import of Figure 7.5.16 is that the output signal to noise variation does not differ from that which would have been achieved given perfect ALC, but for part of the time the output is constant signal and for the remainder constant noise. The assumption has been made here that the detection process is linear down to the lowest level considered. This is not true for envelope detectors as used in AM system, but the principle is the main object here. The effect of a practical detector will be to somewhat increase the noise content at very low carrier levels.

When the other modulations are considered, then similar phenomena occur. For SSB the situation is as depicted in Figure 7.5.16. For FM then the additional factor of operation below threshold must be considered. This can only worsen the noise level for the lowest carrier levels - far more than the situation for real AM detectors.

Before considering this effect in any real detail the action of the mute must be taken into account.

7.5.9 Low Carrier Levels - Mute Effects

Returning to the other factor discussed in sub-sections 7.4.5 and 7.4.4 entails consideration of the effect of mute operation in the quasi synchronous case. For when the resultant carrier level falls below the mute level, then the audio output will be cut off. There will be, therefore, for low level carriers, breaks in the audio signal at the periodic rate of the beat frequency.

Thus there are two thresholds of operation for decreasing input levels where first the ALC/limiting ceases to function and then the mute operates. These can be shown diagrammatically in Figure 7.5.17 which is the phasor diagram of Figure 7.5.2 with the loci of the ALC/limiting and mute thresholds shown and also that of the resultant.

Figure 7.5.17 illustrates, to scale, the case for AM where the mute is set at 1μV (where the signal to noise ratio is to be taken as 12 dB), the ALC threshold is 2μV, the larger carrier P is such that on its own it would result in a signal to noise ratio 10 dB above the mute setting (ie 22 dB) and the smaller carrier Q is 1 dB below P. The audio output of Figure 7.5.16 is then modified and appears as in Figure 7.5.18 where the ordinate is voltage (with the signal arbitrarily set to 1) to allow the zero output condition to be depicted. If the difference in carrier levels were a little greater (3.3 dB) then the mute would not come into action and the output would have the form shown in Figure 7.5.16.

Figure 7.5.17 covers also, in principle the case of SSB: FM will again be similar but with the enhanced noise contribution due to thresholding as discussed previously.

These effects further complicate any attempt to theoretically quantify the audible effect of QS operation. If the carrier difference frequency is very small then during the muted period one or more words may be lost, for greater difference frequencies the loss may be a syllable or two, and at higher difference frequencies the breaks could be at a syllabic rate. In practice the situation is further complicated by the fact that the mute does not act instantaneously but has associated attack and decay time constants, which render the mute inoperative at higher difference frequencies but at lower difference frequencies displaces the cut off period from the time of worst signal to noise ratio (and worst distortion) and in effect cuts off a portion of usable signal.

7.5.10 Performance as a Function of Levels of Carriers

The modulation distortion effects described in sections 7.5.4 to 7.5.7 were primarily a function of carrier level ratios whereas the ALC/limiting and mute threshold effects are dependent on both absolute and relative levels of the carriers. As the carrier levels are increased then it can be seen from Figure 7.5.17 that, for a given ratio of these levels, the portion of time over which the mute operates decreases; whilst for a given carrier level the variation in muted period initially increases for increasing ratio of QS carriers although it will eventually fall for a mute (threshold) setting which is below the strongest carrier.

Thus one measure of performance as a function of input carrier levels could be the fraction of the cycle (or the fraction of total time) for which the mute is operative - the shorter the better. If the mute operation level is taken as a threshold (not the normal FM threshold necessarily) then curves can be drawn for the consequential fractions of time for which operation is below the threshold. The situation is analysed in appendix G and the resulting curves are shown in figures 7.5.19 and 7.5.20.

The former shows the percentage of time for which two carriers will cause the threshold to be broken (mute to operate) as a function of the ratio of the stronger of the two carriers to that threshold. Several curves are shown with the ratio of the two quasi-synchronous carrier amplitudes shown as a parameter. An interesting feature which can be seen from Figure 7.5.20 is that, starting from the point where the carriers are equal in amplitude (carrier ratio = 0dB) the situation in general worsens as one of the carriers is reduced in amplitude, ie the QS effect ,as measured by time below threshold, is greater. Thus at very low carrier levels it appears to be slightly advantageous to have a very strong beat, but this is an extreme condition and ignores mismatched modulations.

7.5.11 Performance for More than Two Carriers

The rough measure of performance deduced in the last sub-section at least has the merit of providing a means of calculating the effect of having more than two carriers present. This situation will occur fairly frequently in a quasi synchronous scheme since there are regions where a mobile will be at the confluence of three stations.

The form of the resulting carrier amplitude for such a three station reception is shown in Figure 7.5.21. On top of the regular two station beat is superposed that of another beat in the form of a carrier amplitude modulation.

It will be seen that the situation is complex. The dashed line in this figure represents the threshold which could be used for performance measurement. The actual conditions are: strongest carrier of amplitude 1 and threshold 10 dB below this, two equal weaker carriers 5 dB weaker than the strongest and offset from the strong carrier by 1 Hz and 0.8 Hz respective. It will be seen that the threshold is cut for some 5% of the time whereas if any combination of the strongest and one of the weaker carriers were present only then the threshold would not be cut - ie the situation has apparently been worsened by the introduction of a third carrier.

The complete situation has been analysed in appendix G and the results are presented in Figures 7.5.22 to 7.5.25. They show the percentage of time for which the combined carriers lie below the threshold, and take the same form as Figures 7.5.19 and 7.5.20. They assume that the third carrier will be equal to or smaller than the second one by the ratio shown on each figure. It can be seen that the time diagram of Figure 7.5.21 corresponds to the peak of the 10 dB parameter curve of Figure 7.5.25.

The, perhaps intuitive, feeling that three combined carriers are always less likely to fall below threshold is not borne out by these figures. Compared to the two carrier case at ratios of major carriers a little above OdB, the addition of a third carrier worsens the percentage of time spent below threshold, and the value of this ratio needed to eliminate QS effects is increased as the level of third carrier (relative to the second) is increased. So that, whereas a particular strong carrier ratio would avoid any occurrence of the below threshold condition for two carriers, the addition of a third can mean that the threshold is crossed. This deleterious effect is however more than compensated for at the positions where the two strong carriers are very nearly equal - at or around OdB. For the two carrier situation this is the critical area and here the addition of a third carrier is always beneficial. This can be seen by comparing Figure 7.5.25 at the position where all three carriers are equal with Figure 7.5.20 where it will be seen that for the case where the threshold level is 10 dB below any one carrier that there has been a reduction of time spent below threshold from 10% to 2%.

Moreover the peak of the curves for any one value of major carrier to threshold ratio, although shifted to the right as the two weakest carriers become more equal, is always reduced in amplitude by such an operation.

In addition to these effects the plot of amplitude of combined carriers against time will not have the beats and threshold crossings in as regular a manner as for the two carrier situation. The depth of amplitude troughs will differ from one to another and the pattern of the beating will repeat at the lowest common multiple of the offset frequencies. This can be seen from figure 7.5.21 and is likely to reduce the impairment.

Thus the two, near equal, carrier situation is more critical than that of three carriers.

7.6 A Consolidation

7.6.1 Summary

Chapter seven opened with the placement of the modulation question in the WARC programme and the position of the author as adviser. Having conducted the foregoing analysis the major points were used as the basis for a presentation on the subject to the rest of the programming committee and other interested parties (including users) this was done against the background of virtually universal predisposition on the committee's part to the use of FM. The true reasons for this were hard to find.

There was a general feeling that since a large portion of UK users used FM and that the rest of the world did so, then it must obviously be the best (particularly since no AM was used in the USA). There was a degree of 'folk lore' concerning the magic of capture effect, signal quality, and 'noise quieting'. The question of power efficiency was raised (ie battery power to RF signal - or more pertinently to output signal to noise ratio); whilst this is of significance in personal radio use, it was of little significance here where the transmitters were either fed from the mains at base station or heavy duty batteries in vehicles - which inherently maintained them in a good state of charge (in any case the major consumption of charge is due to the listening only mode in the vehicle). The difficulty of obtaining semiconductor output stages capable of delivering the AM mobile transmitter power at the upper end of the envisaged frequencies (some 220 MHz) was cited. It was stated that we would have to 'pay-through-the-nose' for our equipment, not only because of its special requirements, but because it was AM - a non standard form for which national and international competition would not act to keep prices down. In fact the author had been specifically told to ignore the economic aspects which would form part of a larger frame-work of consideration and posed much deeper political questions.

There was certainly much pressure from manufacturers with vested interests - particularly those who only produced FM equipment. This will be further commented on later in section 7.7

At the start of the work SSB had not been considered a really serious contender since ther were a large number of practical difficulties and no practical equipment available (beyond some laboratory prototypes) but this modulation had been included in the analysis and there were growing pressures for the Directorate to lead the way in adopting it, and practical - although still not production - equipment became available towards the end of the study. It was obvious that if spectrum congestion was not relieved - particularly for the radio telephony case - then there would be sufficient pressure to ensure that viable SSB equipment was available, given that it could be shown to be, at least theoretically, acceptable. Comparative tests of AM, FM and SSB for the single station case were conducted by the UK regulatory body (Home Office 1981), with the author on the steering committee and these showed that the radimentary SSB trial was at least as acceptable to the general user as 12.5 kHz, FM or AM.

Thus the presentation tended to concentrate on AM vs. FM, with SSB held as a possibility for the future. A summary chart was produced and is reproduced below as table 7.6.1.

Table 7.6.1

Theoretical Modulation Comparison, 12.5 kHz Channelling

Criterion	AM	FM
Single base station operation in general	=	=
Mute Tail	✓	x
Punch through	✓	x
QS Operation:- tolerance to modulation mismatch	x	xx
Compatibility with SSB	✓	x

(✓ = better, x = worse)

The last factor requires some comment. If SSB were to be adopted ultimately, that is not initially for WARC, but some time later, then there would have to be a change-over period when compatibility between the old system and the new was required. This could be easy for AM, SSB sets could be introduced in say the mobiles and their reception of AM would be acceptable. If SSB were transmitted then reeption of this on an AM receiver could be tolerated for a while if the carrier level were held high - it could be reduced later when all SSB operation were possible. On the other hand it is very difficult to see how FM sets could be compatible during a change over situation.

7.6.2 Intermediate Outcome

The commitee did not wish to change its entrenched views on modulation as a result of the presentation. Even the inclusion of recordings of modulation comparison of a bench simulation of quasi-synch operation did not completely dispel their long held views.

Whilst it was agreed by the author that the balance between AM and FM was largely finely balanced, he maintained that compatibility with SSB was a major factor and so was that of tolerance of QS systems to mismatch.

This is worth elaborating. Whatever new system were to be introduced it would have to make use, at least initially, of the existing radio links from control to base station. Some of this equipment was old, but would not be changed, and relatively little attention had been paid to its procurement and subsequent alignment for spaced carrier working let alone QS. But some QS schemes had been implemented with varying degrees of success, depending mainly on the degree to which the modulation could be aligned and maintained. It was known therefore that link alignment would be a burden for the WARC implementation. But it had been shown to be possible for our AM schemes whilst other FM schemes had failed (more of this in 7.7). Indeed previous work under the conduct of the author (Kent 1979) for a QS FM personal radio scheme had shown the extreme intolerance of such a system. Thus even if the costs of the more critical link alignment could be contemplated (new equalising and test equipment) the time and manpower resources would not be available for its initiation let alone its maintenance.

The committee in general agreed to remain sceptical as to the author's findings. He therefore determined to mount a practical demonstration.

Such an outcome was not surprising and had been anticipated. There has been a running thread through the theoretical discussion to the effect that, even if exact and quantitative prediction of such factors as audio harmonic content could be made, there were so many factors each of which could only be assessed for true value on a subjective basis, that the only way to answer the comparison question was subjectively!

7.7 Practical Demonstrations

7.7.1 Preliminary

The first steps in conducting the demonstrations were the setting up of a bench test set to simulate a two station QS scheme. Such a set is shown in Figure 7.7.1.

A VHF receiver (operating around 100 MHz) was fed from the signal generators via a combining hybrid and common attenuator. It was considered advisable to adopt this arrangement since the signal generator levels could be accurately checked at a relatively high level, say 1 mV, by means of a millivoltmeter or power meter, and then both brought down to the μV range required by the receiver in a common attenuator. This avoided any errors in tracking between the two output attenuators of the signal generators. They had a common modulation source which could be speech from a tape recorder or sine wave signal generator. Great attention was paid to the levels of the speech and a professional speaker was used for his consistency of delivery [the tape formed part of a specially purchased audio visual training aid]. The audio signal could be subjected to differential phase shift, differential delay and of course differential amplitude using the input level settings of the signal generators. The differential units are worth special mention:-

i. The differential delay unit consisted of two bucket-brigade analogue shift registers which produced a non frequency conscious delayed version of the analogue input. The delay was dependent on the clock rate so, by driving the two halves at different rates, a controlled differential delay could be produced.

ii. The differential phase shift unit contained two polyphase networks with outputs arranged to be in phase quadrature. By inversion, summing and weighting these outputs, the unit provided adjustable phase shifts of 0 to 360° relative to one output, with the value constant across the audio frequency band.

The frequency shift unit was included to simulate the small frequency shifts which could be experienced if one or more of the links were implemented via a multiplexing unit as might happen if a telephone link were used.

This then was the test set used to provide the tapes for the presentation previously described, since the signal generators and the mobile receiver could be switched between AM and FM modes.

The set amply confirmed the theoretical predictions made in sections 7.4 to 7.5. But it must be admitted that there is a significant difference between, on the one hand, a bench demonstration the parameters of which are static between settings, and which is assessed by an audience in a quiet committee room; compared to listening to a real scheme in a mobile which may be moving, or even if stationary has other vehicles moving around it and an entirely different ambience.

7.7.2 The Demonstration Arrangement and Outcome

The practical demonstrations took the form of simulating a three station county coverage scheme on a reduced scale.

Figure 7.7.2 shows the geographical disposition of the trials test set - termed the Radio Test Area. Three transmitting sites form the apexes of an almost equilateral triangle of side some 20 km. Each site was fed by its own individual link (on its own frequency) from a control at the laboratory. Routes could be taken which covered outer London suburbia through small towns and villages, to rural areas. All classes of roads could be covered including a motorway.

Each base station received the audio from its link and this was used to drive three transmitters - AM, FM and SSB. Each modulation had its own radio frequency and this was common to all three sites. Thus QS operation with two or three transmissions could be simulated for each modulation. The audio signals fed to two of the sites were identical whilst that to the third (Base Station P) could be varied using the arrangement for modulation mismatch described in the last section. The block diagram of the arrangement is shown in figure 7.7.3.

The mobile equipment consisted of a standard emergency services transmitter/receiver which could be switched between AM and FM modes and a second set, kindly loaned by Pye Telecommunications as one of their development/evaluation models. [The SSB transmitters were a last minute addition to the demonstrations and built by P Bridgeman].

The conduct of the demonstrations consisted of taking a few of the interested parties, in the mobile, to a location close to base station Q which simulated strong single station conditions for all modulations. The observers could call for either operator messages or the pre-recorded tape and they could choose which transmission they wished to listen to at any one time by suitably switching the receivers. The transmitters at any one site all radiated simultaneously and this was continuous unless the observers requested that one or more sites be disabled.

Having established the reference 'good' conditions the mobile made a normal journey roughly towards base station P. This took it through weak signal areas (the transmitter powers were equated at each site and held to a low level), two station QS and three station QS.

Following this the vehicle was put in a side road roughly mid way between base stations N and P, at a point where the signals from stations P and Q were nearly equal. Here base station N was switched off and the results of modulation mismatch assessed for the three modulations, again the observers could call for any conditions (except variations in carrier offset frequency).

Finally a fast run could be made on the motorway to demonstrate the effects of the so called 'doppler shift' or travelling through a weak fading field which showed the effects of FM noise bursts.

The outcome of the demonstration was to convince all (although some still claimed reservations) as to the accuracy of the theoretical predictions and that AM was preferred to FM. The SSB system was judged to be full of potential but not preferable to AM - this was not surprising in view of its undeveloped nature (straight pilot carrier SSB with carrier 10 db down on peak modulation and 5 kHz channelling filters).

Some recordings made on this system are available with this publication, on cassette tape.

7.7.3 Additional

Thus the technical decision could be taken, on both theoretical and subjective bases, that the modulation which should be used post WARC would be amplitude modulation. There were however many other factors to be considered many of which are not really pertinent to this narrative. For the sake of keeping the record and for completeness it is however worth relating briefly some of the other factors which involved the author.

It has been said previously that there was pressure from manufacturers to adopt their favoured modulation - not surprisingly. Requests were therefore made, as part of a wide range of consultations with the manufacturers, to be given the opportunity to hear their quasi-synchronous schemes. Only one in the end obliged, and that in a very open manner, whilst the others managed to avoid the situation. The result of this was to further confirm the choice of AM.

One other of the emergency services had long operated a FM QS scheme (one not taking its services from the Directorate), and we liaised on experiments and demonstrations. Their current scheme benefited from a number of advantageous factors. These were; 25 kHz channelling, favourable geographical disposition from the frequency reuse viewpoint, favourable topography which admitted very dominent base stations on high hills to cover the bulk of the county on a coastal plain, the use of high transmitter powers, the consequential very high signal strengths, more staff for scheme maintenance than that used by the Directorate, time and ability to optimise the system. Thus the present system worked very well and was highly regarded locally but it was difficult to see how this could be repeated nationwide in the WARC scenario. Their experiments with AM QS could only be conducted as a side-line and were open to criticism, certainly on a purist experimental basis. The outcome from this was not to change our decision; in their case, they at one time decided on AM and then reverted to FM. It will be very interesting to compare performances when the WARC systems have been installed.

7.8 AM Alignment Parameters

7.8.1 Introduction

The decision having been taken that amplitude modulation be used for the post WARC mobile system, it was obviously necessary to determine the tolerance of an AM QS scheme to the various factors which would cause it impairments. From these overall limits the specification and maintenance figures could be derived for each portion of the contributory causes - mainly the fixed links.

Such limits, it had been shown, would need to be determined on a subjective basis, and to obtain representative results thus meant a listening panel. Thus an experiment was envisaged in which many recordings of an AM QS system were made for every possible combination of factors - obviously an impossible task even before considering the reaction of the listening panel to such an exercise. Furthermore it had been shown that such an experiment would have to very closely follow the conditions met with by a mobile in a practical situation - hence the most sensible way was to conduct the recordings using the radio test area, already described, and a moving mobile.

But two important preliminary steps were taken. The first of these was to use the theoretical analysis described in section 7.5 to point to the most critical parameters and the likely range of parameter limits. The most important factors from this were:-

- misalignment distortions most noticeable when carriers are nearly equal and in antiphase

- alignment of modulation indexes less critical than alignment of modulation phase

- mute and ALC effects most noticeable when carriers are near equal and at low level

- frequency difference (the offset) producing beating effects which are worst at low and high values.

The second important step was to conduct a preliminary investigation using the bench test set-up described in section 7.7 for the initial comparisons.

7.8.2 The Laboratory Simulations

In order to plan for the most realistic assessments of a QS system as a field trial, initially a two station scheme was simulated on the bench. This served to confirm the theoretical guides, provide an indication of what values the parameters should range over, and trial the method of assessing allowable parameter limits. Whilst admitting easy control of one parameter at a time (or several in combination) it could not simulate a moving mobile experiencing not only QS effect but multipath and range variations.

This phase (subsections 7.8.2 and 7.8.3) of the work has been reported on under the name of one of the author's assistants (Asque 1982) but as said earlier will be included here for completeness. The equipment arrangement was that of Figure 7.7.1 with only the AM portion in use.

Great care was taken to ensure that the two parts of the system were accurately aligned and that the parameters could be calibrated. To this end peak programme meters were used for the measurement of the audio tone and the speech levels from the tape recorder (this had been recorded by a professional speaker); the signal generator frequency synthesisers were driven from a common high stability source so that frequency offsets of hundreds of Hertz down to zero could be achieved, and two attenuators external to those in the generators were used so that both absolute and relative levels could be set accurately.

Initial tests were conducted to see if any correlation could be established between an objective measurement of the receiver's audio output and the subjective one as judged by the experimenters. The objective measure considered most likely to achieve this was SINAD (Signal In Noise And Distortion) which could deal with the distortion of a modulation signal and its noise content. It was quickly established that no correspondence could be found, so the audio output was recorded for later assessment by a panel of listeners.

Details of the assessments are given by Asque, briefly the panel were asked to rate the channel for Readability and independently for Impairment according to the five point scale shown in Table 7.8.1 which corresponds to CCIR recommendations (CCIR 1978a).

TABLE 7.8.1

Scoring Definitions

READABILITY	IMPAIRMENT
5. Perfectly readable	5. Imperceptible
4. Readable with slight difficulty	4. Perceptible but not annoying
3. Readable with considerable difficulty	3. Slightly annoying
2. Barely readable - some words distinguishable	2. Annoying
1. Unreadable	1. Very annoying

Each session commenced with a briefing for the listening panel on the procedure for the tests and a reference signal was played. Although the ambience of a mobile could not be reproduced care was taken to simulate the response of a mobile radio by using a loudspeaker from such a radio. Table 7.8.2 shows the parameters which were varied. The panel were given time to assess each level of variation of each parameter which was presented to them in random order.

TABLE 7.8.2

Parameters Explored for Bench Tests

1. Carrier frequency difference
2. Modulation index difference
3. Differential baseband delay
4. Differential baseband phase shift
5. Differential baseband frequency shift
6. Absolute and relative carrier levels
7. Receiver mute setting
8. Receiver local oscillator drift.

Figures 7.8.1 and 7.8.2 present a selection of the results showing graphically the mean scores awarded for the two criteria.

7.8.3 Discussion of Results of Laboratory Simulation

The results confirmed the crticial areas deduced from the theoretical analysis and limited the area which needed exploration in the field trials. It was shown that no baseband frequency shift could be tolerated, whereas local oscillator frequency drift was of no practical significance - up to 3 kHz could be tolerated for the 12.5 kHz channelled receiver under test.

Any carrier frequency separation was found to cause impairment of the demodulated signal by introducing a beat note with a frequency equal to the carrier difference frequency. Up to about 5 Hz difference the impairment was perceived as an amplitude variation, between 5 Hz and some 30 Hz a tremulo or warble was heard, above 30 Hz the distortion has the quality of a superimposed tone. The subjective assessments showed that a small frequency difference was preferred with an optimum value in the range 0.5 to 2 Hz. As expected below 0.5 complete words could be lost whilst above 2 Hz readability was impaired. For the field trials therefore, where it would be very difficult to accurately control variable frequency offsets, a fixed value nominaly set to 1 Hz was chosen, this also duplicates the conditions which could realistically be maintained in practice. The improved scores for offsets above 4 Hz shown in Figure 7.8.1 applied only when there were no other distorting factors present (eg phase/delay).

As expected the scores given decreased as either the audio phase difference or audio delay difference was increased. There is a tendency of measuring, for spaced carrier systems, the audio phase difference at one frequency only and compensating for this by incorporating in the appropriate path a delay unit which gives an equivalent phase shift at this frequency; this assumes that there is a linear change of phase with frequency over the whole frequency band. Measurement of the complete characteristics of such paths has shown that this assumption is not in general valid, and it should be noted that the results shown here apply to either constant phase shift across the audio band or constant delay (linear phase) across the band.

The composite effects of increasing delay when there was already a phase shift is shown. The apparent anomaly of initially increasing scores can be accounted for by the fact that in the test configuration used for the bench tests the sense of increasing phase was opposite to that of increasing delay.

7.8.4 Field Assessments

The bench tests were followed by making recordings, in a mobile, of conditions actually experienced in a simulated county coverage scheme. This consisted of three base stations located as shown in Figure 7.7.2 and connected as shown in Figure 7.8.3 all radiating to the mobile on a frequency close to 100 MHz. Again all paths were checked for accurate alignment, and the phase and delay units were arranged to act in the same sense. By using separate fixed link transmitters at control each on its own frequency, one base transmitter of the three could have the desired baseband variations imposed upon it. In this respect the scheme differed from that of a normal county coverage one which uses a common outgoing link transmitter.

Only a limited number of tests were considered necessary to assess the effect of:- mobile motion, multipath, mutiple transmitters, ignition noise, and local environmental factors such as buildings, other vehicles etc. The range explored is shown in Table 7.8.3. The opportunity was taken also to assess the effects of diversity reception using a two aerial commutating system (Fudge 1978). Overall some 90 recordings were made for various combinations of the factors shown in Table 7.8.3 and these were played to the assessing panel in the same manner as those of the laboratory tests.

TABLE 7.8.3

Parameters Explored for Field Trial

OPERATIONAL	CONTROLLABLE
1. Moving mobile	1. Modulation index difference
2. Stationary mobile	2. Differential baseband delay
3. Multiple transmitters	3. Differential baseband phase shift
4. Diversity reception	4. Absolute and relative carrier levels
	5. Receiver mute setting.

7.8.5 Overall Results - Discussion

It was quickly confirmed that QS operation was only critical in areas where comparable signal strengths were received from two transmitters at a low level. To be of significance when in motion this equal signal area had to be maintained over the run of the mobile. When moving, the 'fast fading' effects of multipath were superimposed on those of QS field variations. Only under the conditions of low and nearly equal carrier levels without significant multipath could the so called QS Doppler Effect be noticed. The QS Doppler Effect can be considered to result when a mobile has a component of motion towards a transmitter which causes the received carrier's frequency to increase by virtue of the Doppler Effect. Thus the difference or beat frequency between the carriers will change. If the mobile is moving away from one transmitter and towards the other the effect will be greater and will appear as a Doppler induced addition (positive or negative depending on initial factors) to the QS offset frequency. An alternatve view of the phenomena is to consider the mobile moving through the interference field pattern of the two transmitters. On this basis the requirement that the two field strengths be nearly equal over the length of the run surveyed for the phenomena to be apparent is obvious.

Diversity reception, with the commutator (switching at a rate of 7.5 kHz), for a moving mobile was found to give no significant improvement in audibility, whereas the introduction of measurable level of a third transmitter was always found to be beneficial. Movement in general was found to be beneficial due presumably to the factors discussed above concerning the QS Doppler Effect. For the mobile tests the vehicle was driven at a constant 40 mph in the same direction over a well defined short route with very nearly equal signal strengths from two transmitters. The subjective assessments for this were found to cover a wide range of scores. Further analysis showed that some runs benefited from the occurrence of carrier nulls coinciding with natural gaps in the message, whereas in other runs the reverse was the case. A judicious allowance for this factor was therefore made in assessing the recommended parameter limits.

The most critical situation was that of a stationary mobile in a low equal signal area - virtually the situation which could be simulated in the laboratory.

Figure 7.8.4 shows in graphical form some of the results obtained for the composite field trial operational factors. Unless stated otherwise the conditions were: phase shift 0°, delay difference 0, modulation index 80%, received signal levels 8 Vemf, carrier offset 1 Hz nominal. It will be seen that the results tend to follow those of the laboratory simulation and in view of the discussion above this is not surprising. In particular there is no sharp degradation in score for any variable. Thus, in the absence of a well defined parameter threshold, the limit which should be prescribed for any parameter is not crticial within small variations. However quantitative measures can be derived from the compendium of results obtained from the trials bearing in mind that many of the parameters are likely to be near their limits at the same time.

7.8.6 Recommended System Parameter Limits

From the tests conducted the following parameter limits are recommended for the design and maintenance of Quasi-Synchronous AM systems operating at VHF. The differences are those between adjacent base stations.

1.	Carrier frequency difference	0.5 to 2 Hz
2.	Modulation index difference	less than 6 dB
3.	Differential baseband delay	less than 100 μs
4.	Differential baseband phase shift	less than 60°
5.	Differential baseband frequency shift	Zero
6.	Absolute and relative carrier levels	

When the signal from the dominant transmitter gives an input to a mobile of less than 10 μV emf, then any other carrier should be at least 6 dB below that.

7. Receiver mute setting

A setting that just mutes the receiver with no carrier present will not significantly affect the readability of any received QSAM signal.

8. Receiver local oscillator drift less than 3 kHz.

Notes

i. The total differential audio phase shift at any frequency should be less than 60^{o} or $(0.036xf)^{o}$ whichever is the greater. (Where f is the baseband frequency in Hz).

ii. The levels given in 6 above may be dependent on the performance of the receiver. The one used for the tests was representative of those in use by the emergency services but which is now of obsolete design. For improved receivers these levels may well be lower.

8 Propagation Factors III

8.1 Frequency Bands Considered for Operations

8.1.1 The Situation

A major task for the WARC planning was that of:- seeking new frequency bands, assessing them, and then acquiring them. It was apparently no function of the regulatory body - who in effect decided that the the UK emergency services would not in future be allowed to use the 100-104 MHz band - to simply find an alternative. Since all suitable bands were in use then it would entail displacing an existing user from his possession. To the extent that the usage of the suitable frequency bands was more intense in some areas, and heavily constrained in others, then they were able to offer guidance as to likely potential regions and provide backing and support to the Directorate for the negotiation to acquire new bands. The degree of backing and support would of course be dependent on their approval of the Directorates choice and their possible long term plans for overall spectrum usage.

The regulatory authorities suggested that either the bands known as Band III or, for preference, Band I would be suitable for our use. these were broadcasting bands currently in use by the ITV and BBC respectively for monochrome TV broadcasts on the 405 line standard. The abandonment of that standard had long been known and a date had been set for the cessation of broadcasting in these bands; if adhered to this would be compatible with the time frame for change over.

In addition to these regions the Directorate of course retained the 80-84 MHz band and those currently in use for fixed links of 146-148 and 154-156 MHz.

It was necessary therefore to examine the suitability of those bands to accommodate the needs of the emergency services and the author contributed to this assessment to the extent described in the subsequent sections of this chapter.

8.1.2 Some General Considerations

Chapter 4 (Propagation Factors I) derived the simple expression for the received field strength at a mobile and the power input to a mobile receiver. It was shown that the latter was independent of frequency of operation provided that the same types of aerials were used for transmitting and receiving (eg always matched dipoles). From this simple conclusion it would seem therefore that all possible frequencies from a few kHz up to and including microwaves are likely to be equal in performance. This of course ignores other relevant factors which will be discussed here.

The first of these is the practicality of the aerials. For efficiency the mobile element should be some half wavelength long; now aerials longer than 1 metre become unwieldy, and over two meters cannot really be considered practical. So that the lowest frequencies which can be used are 37.5 MHz and preferably 75 MHz. At the upper end the manufacturing precision required increases with frequency, and more particularly the care of fitting the aerial onto the mobile, matching the feeder and the feeder losses all increase. This tends to limit minimum practical aerial lengths to some 15 cm or preferably 30 cm - ie 2 GHz or preferably 1GHz.

Another factor limiting the upper frequency end of operation is that of fast fading. This was described in section 6.1 as a spatial variation of the field intensity, which to a moving mobile, appears as a time variation of carrier strength. If the rate of variation becomes comparable with the lowest baseband frequencies to be transmitted, then there will be corruption and distortion. This will be true for AM due to the inability of the ALC to differentiate between the frequencies, and also for FM since the field strength variations imply also phase/frequency variations. A simple derivation of the upper limits due to this can be made by the fairly crude approximtion to a field variation caused by two equal components travelling in opposite directions (one, say a direct wave, in the direction of motion, and the other, reflected wave, in the reverse direction). Thus the mobile is travelling through a standing wave pattern with a pitch of half a wavelength $\frac{\lambda}{2}$ .If v is the vehicle velocity in m/s then the rate of variation is $\frac{2v}{\lambda}$ Hz. Thus for an upper limit of say 300 Hz and a speed of 70 mph (112 km/h = 31 m/s) λ becomes 0.2m or an upper frequency of 1.5

GHz. To allow some margin here it usual to take the upper frequency for mobile operation as 1GHz.

Although chapter 4 indicated that propagation was independent of frequency, it assumed that there would be no absorption of the RF signals. In fact there is increasing absorption above some 200 MHz by foliage - particularly if wet. This does not seem to have been systematically studied (probably because these bands have been used mainly for broadcasting to a user who is assumed to have a high and directional aerial). It has however been reported as a by-product of several studies.

There is in addition the question of availability of suitable equipment and devices for the new bands. Certainly no commerical equipment available in the UK could be considered for operation above 200 MHz (by the time of writing the 900 MHz bands had been opened to cellular radio and equipment development for that was underway), and there were questions as to the availability of devices for transmitting moderate powers on AM above 220 MHz.

Thus the range over which a search for available bands of operation should be conducted is that of some 50 MHz to 1 GHz with maximum interest in what can be considered to be the prime area for mobile operations of 100-200 MHz. This then fitted well with the bands described in subsection 8.1.1, and each will be described in more detail in the following sections.

8.1.3 Disposition of Go and Return Bands

The use by the Directorate of two frequency simplex, or half duplex, working has already been discussed, and such operation leads to the use of one group or band of frequencies for the outgoing (from base station) path and another group or band for the return channels. The relationship between, or spacing of, these bands is a factor to be considered when choosing new operating frequencies.

The base station can, and probably will wish to (see chapter 10), use separate aerials for the two directions of transmission. Each aerial can therefore be optimised for its band of operation with little or no consideration being given to the position of the other band from the aerial viewpoint. On the other hand it is highly desirable that the mobile uses only one aerial and this therefore must operate well on both bands.

The aerial in common use for the mobile is the quarter wavelength whip and this can be considered to work well over a bandwidth of some 10% of its centre frequency. This then imposes some constraints in the separation of the frequency bands.

As an aside the present operating frequencies of 80-84 and 100-104 MHz are not ideal in this respect. Practice is to cut the aerial (tune it) for optimum results in the 80-84 MHz band so that the mobile transmitter sees a minimum of mismatch and wastes as little of its power as possible. The performance at 100MHz is now much less than optimum on two grounds. Firstly, since it is non resonant at this frequency the conversion of RF field strength to emf signal is poor. Secondly the aerial is mismatched to the feeder, in both resistive and reactive components, and this will combine with the inherent mismatch of the receiver to turn the feeder into a lossy filter whose performance can vary from installation to installation. In practice an allowance of some 6dB is deemed to be made for this aspect and the base station transmitter power increased accordingly. [This aspect raises interesting questions as to what aerial length should be used when making measurements in receive only situations, such as field strength plotting].

Countering the desire for the bands to be close in frequency for the aerial performance point of view, is the need for the base station to operate in a duplex mode (it might be at times desirable in the future that the mobiles could also - particularly for digital data transmission), which will require filters to separate the go and return frequencies. The details of this will be discussed in chapter 10, and that chapter will also start the discussion on intermodulation aspects, to be developed in chapter 11, which also calls for a wide separation between the bands from the filter point of veiw. It was decided early that the separation between bands must be at least that of the width of the bands in order that the dominant 3rd order intermodulation products could be completely avoided, it was expected that this would require filters with at least high performance, but probably realisable ones. Wider separation of two or three time the width of the band would be preferable.

To gain some quantitative feel for the figures involved, assume that each direction of transmissions is 4 MHz wide as before, then the total bandwidth covered for the minimum separation would be 12 MHz. If this were located around a 200 MHz centre, then the fractional bandwidth of an aerial would be only some 6%, whereas if located around 50 MHz it would be 24%. Obviously the higher frequencies would be better from the aerial view point but this would in time make filter realisation more difficult.

8.2 Broadcast Band I

8.2.1 Basic Factors

As described earlier the band extending from 39.5 to 68 MHz was in use by the BBC to convey monochrome TV on a 405 line standard, but the standard was due for abolition thus freeing the band. It was recognised by the regulatory authorities that the upper portion of the band would be of greatest interest due to the increase in background noise levels at lower frequencies, the aerial problem already discussed, and the continuing use of this band outside UK.

The distribution of noise across the RF spectrum for mobile users had been studied as an extension of a Directorate funded University contract. This had initially set out to characterise the noise experienced by a mobile under various operational conditions. The noise was that external to the vehicle and, since it was dominant in urban and industrial areas, was termed man-made (as opposed to naturally occurring effects such as thermal ones). A major component was ignition noise of other vehicles. The aim of the project was to characterise the noise in a form suitable for the designers of a digital data system. The initial work at 100 MHz was extended above and below this frequency and a summary of the findings is shown in figure 8.2.1 as a graph of one of the noise measurement parameters against frequency. This shows the marked disadvantage, from this viewpoint, of operating at the lower frequencies.

Turning to the issue of non UK use of the frequency band brings in the French and Irish dimensions. Those countries have no firm policy of ceasing broadcasts in this band, and the implications were that, even if present services were withdrawn, then others would replace them. Thus it had to be assumed that high power TV transmissions would be radiated from dominant sites in those countries. These transmissions would be received either by a mobile or, at a higher level, by a base station - certainly if operating in coastal counties, and for quite a way inland. The rationale advanced by the Regulatory authorities to admit or overcome this, was to assume that the TV transmissions would be highly structured in a spectral sense. Due to the dominance of certain frequencies (carrier, line synch, frame synch, sound carrier, etc), there should be room to allocate the Directorate's channels in the gaps between these. Thus they envisaged a larger spectral block allocation than would be otherwise required and extended this above the normal broadcast band to give additional flexibility.

8.2.2 Sporadic E Effects

The reception of the Irish. French, and other near continental, transmissions would be function of normal propagation. But anomolous propagation conditions also had to be considered, and here that known as sporadic E becomes of interest. The mechanism of propagation can briefly be described as one of reflection of RF energy from 'clouds' of ionisation in the E layer of the ionosphere. The clouds are some 100 km across and are thought to drift. When one of them is located at the mid-point between a transmitter and receiver, then a portion of the transmitters energy is reflected and received at a beyond the horizon distance. The distance for such communication (or rather interference since it cannot be relied on) is some 1000 - 2000 km. Thus the UK would be vulnerable to transmissions from Southern Spain, Northern Italy, Romania, and the whole of Eastern Europe. The latter could be a particular problem since they appeared to use very high power transmitters and not to conform to the strict limits of the band.

At first the phenomenon was thought to be of a minor nature since its effects were reputed to fall off rapidly at the top end of the band, and the frequency of occurrance was though to be small. It was necessary to confirm and quantify this however.

The first point of note on examining the publications on the subject was that only first order time statistics were available, ie only the average percentage of time a particular level was exceeded for say a year. We needed to know in addition to this the duration of occurrances. By way of example consider a level which might be exceeded for not more than 1% of the time. Even if this was 36 seconds continuously in every hour it would be bad, but if it occurred in greater bunching such as 15 minutes per day, or even 3.65 days per year. then it would be disasterous. Secondly the evidence to support the assertions that the effect was very small around 70 MHz did not seem strong.

It transpired that the University College of Wales (Aberystwyth) had been collecting data for many years on this subject but had not analysed it. So the Regulatory Authorities commissioned an analysis and the author was consultant to this. From the chronological aspects of this study even the early results were not available when the choice on frequency bands had to be made, although apprehension was increasing. A summary of the findings is given here however for completion and continuity. This is shown in figures 8.2.2 and 8.2.3 which are derived from the University's report.

Some comments on this are required. The extent of occurrances at any one frequency measured by the University (the points on figure 8.2.2) were dependant on the number of distant transmitters, on their frequency, their powers, and their angular distribution from Aberystwyth. In each case, the larger the number the greater the effect; there was no firm way of knowing the quantities for these parameters for each frequency to be considered - almost certainly they were different. In any case had these values been known then there was no known way of applying a correction. So the points shown on the graph apply to different sets, and should not be connected. However since they are actual measured results for the existing distribution of transmitters they can be used for planning purposes, but interpolating between the frequencies of measurement (as shown by the dotted line) is fraught - there is however nothing better. The apprehension about Sporadic E still being significant at 70 MHz is confirmed.

The time statistics are also very worrying, there is a significant likelihood of communications being unusable for unaccepetably long periods of time. It is not clear to what extent avoiding the vision carrier frequencies which were used to derive the statistics, would reduce the problem nor whether changing frequencies by several mobile channels (but only a fraction of the TV bandwidth) would change the region of operation.

8.3 Broadcast Band III

8.3.1 Basic Factors

This band, extending from 176 to 224 MHz, formed the second tranche of spectrum which would be freed on abolition of the 405 line monochrome TV standard. This region could be of greater interest to users of mobile radio. It is well within the limits described in sub-section 8.1.2. The background noise levels are significantly lower than for the case of Band I, and there will be no anomolous propagation effects arising from the Sporadic E phenomenon. Some incidence of other forms of anomalous propagation might however the expected - although right towards the bottom end of the frequency range, ducting could be troublesome. As in the case of sporadic E the effect of such anomalous propagation would depend not only in the severity of the phenomenon itself but also onn the intensity and spread of the potential interfering transmitters.

Although section 8.1, following from chapter 4 (Propagation Factors I), said that the signal received by a mobile would be largely independent of frequency (for the same transmit powers and the same aerial types) in fact there will be a change in the nature of propagation with frequency apart from the already discussed ones of foliage attenuation and fast fading. This difference is not manifest in the derivation of chapter 4 since it assumes line-of-sight propagation, but is discussed in chapter 6 (Propagation Factors II). In that chapter, notably section 6.1, it appears as the shadowing variability which is caused by obstruction of the line-of-sight by hills, buildings etc. Under these conditions the RF signal is diffracted around the obstacle and there are methods for calculating the effective loss of signal caused by the obstruction. This effect varies with frequency, the loss being greater at higher frequencies; it can be viewed as a 'sharpening' of the shadow cast by the obstruction. The variablity due to this appears to be still log-normally distributed as discussed in chapter 6 but the mean level decreases with increasing frequency.

Since county-wide coverage is far from universally line-of-sight then there would be a reduction in total cover due to this cause. It could be overcome by an increase in transmitter power or by increasing the number of base stations. The former is to be avoided if possible since it could cause interference to co-channel users (the situation to be considered now is one of inherent terrain variations as opposed to the smooth earth case of chapter 5 on area coverage and frequency reuse). It would also exacerbate the intermodulation effects to be described in chapter 11. The second alternative of increasing the number of base station sites, is easily admissable from the technical standpoint, since area coverage by quasi-synchronous AM has been adopted - increasing the transmitters in such a multi-transmitter mode can only be beneficial. This ignores however the enormous problems of obtaining suitable sites and the associated planning permission for this use, to say nothing of the long time scales attendant on such an exercise.

The factors discussed above could all be quantified and assessed but another, dominant, consideration makes such an undertaking irrelevant.

8.3.2 Other Users

The situation for Band III reflects that of Band I in that, whilst free for assignment to mobile use in the UK, it will still be used for Broadcasting by other countries. The future plans of our geographically nearest neighbours are more firm for this band however. They envisage the continuation, or often the enhancement by very high power TV transmitters, of existing services. Here again the potential for interference to UK mobile users from these Continental and Irish transmissions, due to normal propagation, is severe.

Two situations were analysed:- reception at a mobile, and reception at a base station. The mobile presents the most tolerant situation since its aerial is at low elevation and thus will be a poorer receptor than the normal roof top mounted TV viewing aerial. It could also be considered to be frequently shielded from direct, line-of-sight, reception from the distant TV transmitter. But this cannot be relied on - particularly when the high values for coverage area and percentage of time operation are considered. After making suitable allowance for the probable difference in polarization of the TV and mobile situations, maps were drawn of interference contours from the distant transmitters. This was carried out by the Frequency and site Planning Group of the Directorate, and showed that most of England and Wales would be blighted by this consideration.

At any one blighted location not all the frequencies in the Band would be unusable - depending on the channel of operation of the TV transmission causing the interference. Thus it would be possible in principle to find admissible frequency bands for any one location but this band could not be used universally in the UK. To make use of this it would then be necessary to have a number of frequency bands for emergency service use in UK - a situation which would cause many opertional, logistic, and financial problems.

The use of the Band III for mobile operation was advocated by the Merriman Report (1983) which was published subsequent to this investigation. It suggested that Band III could be used as just described by allotting certain frequencies in certain regions. This might be viable for users with no more than a regional interest. The findings were based on a somewhat less onerous basis than that conducted by the Directorate and did not apparently take account of manufacturer views on the viability of such a spread of total bandwidth.

When reception at a base station is considered then this is less tolerant than that of the mobile. Aerials are invariably mounted on high towers and the towers located on dominant hills. Thus 'good' reception of distinct co-channel TV transmissions is to be expected and the area so blighted will be greater than that for the mobile reception case.

'Split band' working was assessed in which the outgoing frequencies were located in Band III and the return frequencies remained in the 80-84 MHz band. But the problems of mobile aerials, the cost of split band sets, and the general uncertainties of 'security of tenure' in the TV bands all indicated that this was not viable.

8.4 VHF Mid Band

8.4.1 Introduction

The third main option lay in making some use of the frequency bands of 146-148 and 154-156 MHz which were presently in use for the fixed link services.

These bands were already considered to be overcrowded and a policy of reducing channel bandwidths from 25 kHz (in an FM mode) to 12.5 kHz had been declared. [It is interesting to note that no cognisance had been taken of the effects of the necessary deviation reduction on link performance as discussed in chapter 7(7.3.2) and appendix E]. Even a full implementation of that policy, and the complete displacement of the whole of the channels assigned to a particular Directorate user, would yield only one free MHz in each of the bands. Even if all the fixed links were removed as well, then the total available was 2 + 2 MHz as against the 4 + 4 desired, and the complete removal of the fixed links within the WARC time scales, let alone by the start of the main frequency change. could not be considered viable from the sheer logistics point of veiw - neglecting the problems of relocatingthose services elsewhere in the spectrum.

The prospects for the use of this band were then not bright but at least 1 + 1 MHz of usuable spectrum could be realised and it might possible to build from this by acquiring more spectrum from adjacent frequencies. This would at least prepare the ground for capitalising on the eventual removal of the fixed links. [There had been thoughts for some time on the desirablity of implementing the links at microwave frequencies since the Regulatory Authorities no longer approved of VHF fixed links, and multiple carriers over the same path could be carried on one multiplexed microwave carrier]. Thus the surrounding spectrum was surveyed to see if any suitable portions could be acquired, the disposition of use is shown in figure 8.4.1.

8.4.2 Details of the Band

The first point of note from figure 8.4.1 was that two of the portions of spectrum adjacent to the Directorate's bands were assigned on an international basis, 144-146 MHz to Amateur use, and 154+ for Maritime purposes. Obviously the users of the latter of these could not be displaced without extensive international agreement which would taken many years, even if feasible in the first place, and even then there would have to be a lengthy period of grace for the equipment to become obsolete. Therefore no inroads into that band could be contemplated. Similarily the Amateur band was assigned internationally and used as such. Thus even if it were contemplated to displace UK amateurs from this band, then there would be interference on a two-way basis between continental amateurs and the UK emergency services; therefore this band also was unavailable.

This immediately limits the room for manoeuvre since the upper band can only be extended downwards and the lower band upwards. If each were made 4 MHz wide then there would be only a 2 MHz gap between them which by previous reasoning is insufficient. The best that could be contemplated on this basis would be to acquire one extra MHz on each of the existing bands ie operate 146-149 and 153-156 MHz. The gap between the bands would then be just sufficient for the minimum intermodulation requirement.

At the time these considerations were being made the 3rd order intermodulation effect was known to be a likely problem and to be avoided. This was due to the coupling between base station transmitters. The author and his team were very apprehensive about the need to avoid higher order products emanating from that cause and the site non-linearities (these are discussed firstly in chapter 10 on Base Station Design, and more fully in chapter 11 on Intermodulation). However no quantitative values were available at that time so that the case for needing to consider such effects was judged by others to be unproven and therefore neglected.

The band from 148 MHz upwards was in use for military purposes and negotiations were put in hand by the Frequency and Site Planning Group to acquire it, but the band 153-154, as can be seen from figure 8.4.1, was in use for fairly recently assigned purposes. The upper half of this 1MHz was used for Radiopaging on a nationwide basis, and there was a high level of deployment of equipment - mainly the paging personal sets. Whilst perhaps not expensive on an individual basis the problems involved in recalling them can be envisaged. There was also the point that some of the users in this band had been fairly recently displaced from the 27 MHz paging band. Therefore the reassignment of this half MHz to the Directorate could at best be viewed as a long term exericse - much longer term than could be tolerated for the WARC frequency change. It seemed possible however that some assignments could be made in the lower half MHz ie 153-153.5 MHz although these were likely to be specific channels rather than a complete and contiguous block.

Therefore the possiblity of acquiring some of the band 150 - 153 MHz was explored. This it will be seen was in use by the Radio Astronomers. Only two sites in UK were thought to be involved in the use of this band; Jodrell Bank and Cambridge, for the purposes of receiving only. It transpired that their work could be accomplished with significantly less than 3MHz, although it was recognised that the very high sensitivity of their receivers would make an effective guard band desirable between the bands they used and any transmitters in nearby bands. But if more than 1 MHz were to be acquired it was doubtful whether it could effectively be used due to the fact that the upper band would now spread from 152 to 156 MHz ie 4 MHz and that the 3rd order intermodulation products would extend down to 148 MHz. [The gap in the 152-156 band would have no significant effect as far as the spread of intermods was concerned]. Thus it would be only sensible to acquire the one MHz of 152-153 MHz of the Radio Astronomy band and this would be at the expense of the use of the one MHz of 148-149 MHz.

Therefore perhaps three MHz total might ultimately be available in the upper band, 152-153 and 154-156 MHz. To pair this with a lower band of equal total width, an extra 1 MHz below the 144 - 146 MHz amateur gap was sought. After negotiations with the present user, the one MHz of 142-143 MHz was obtained.

8.5 Outcome

8.5.1 Consolidation

The use of three possible regions of the spectrum has been discussed in the previous sections. The regions were Broadcast Band I, Broadcast Band III, VHF mid band.

The major problem of using either of the first two was found to be the interference from Irish and Continental broadcasters who would still be using them for broadcast purposes. When WARC reassigned the band 100 - 108 MHz to broadcasting, and thus occasioned the whole of this exercise. they also re-designated the broadcast bands I and III to a dual role of broadcast and mobile use. But the broadcast mode was 'primary'. Thus it could at any time override and displace any mobile assignment, furthermore although the UK designated it as a mobile band it could not protect its users from external broadcast interference since this was its 'primary' designation.

If the Directorate were to find suitable portions of either band, then they would be vulnerable to another country starting broadcasting services in the region of frequency so chosen. This could happen at any time and could even occur due to a change of heart by the UK regulatory body (who would have eventually to bow to any political desires in this respect). Therefore the Directorate would have no real security of tenure and would be in exactly the same position of being a 'lodger' in broadcasting bands as it was at 100 - 104 MHz. The saga could then start all over again in the future.

For these reasons the attention became concentrated on the VHF mid band solution. But section 8.4 has shown that there would be great difficulties of use there. The existing fixed links would still have to be accommodation in these bands, at least for part of the WARC change over and the consequential increase in the number of transmitters close in frequency (fixed links and main) would cause site problems - particularly intermodualtion products. Furthermore only some 3 + 3 MHz including the fixed links was likely to be usable in place of the 4 + 4 MHz desired.

8.5.2 Final Form

Whilst rejecting the use of broadcast bands it was stated that some spectrum above the normal band I range could be available. This would not be used by our near neighbours but was in use in Eastern Europe for broadcasting TV and could therefore suffer from the sporadic E effects (at that time still to be quantified). The use of a portion of spectrum here was however attractive since it could be paired with the 80 - 84 MHz portion, so the 70.5 - 71.5 block was acquired.

Since the emergency services channels were to be 12.5 KHz spacing then there were some grounds for assuming that less than 4 + 4 MHz would be needed for mobile use. But certainly more than 2 + 2 MHz seemed to be indicated since a halving of channel bandwidth had to be undertaken even before the onset of WARC in order to cater for growth - and this was a factor which obviously needed to be planned for at this stage. Therefore if the fixed links took 1 + 1 MHz of the 3 + 3 MHz available in VHF mid band the 2 + 2 MHz left would not be enough and some users would have to use the VHF low band (70/80 MHz).

After extensive enquiries it was found that there was virtually zero interoperation between police and fire users on the mobile radio interface. All desired communication was achieved via a land-line connection between their headquarters. It was therefore decided to allocate the 70.5 - 71.5 MHz and a portion of the 80 -84 MHz bands to fire service mobiles only. A 1 MHz band could be considered generous for this user but the ability to be flexible in assignments (even dynamic) was deemed necessary to overcome possible sporadic E effects. Thus the VHF mid band could be assigned to fixed links (police and fire) and police mobile channels.

In each case the direction of working, ie which band should be used for base station transmit and which for receive, was effectively pre-determined. Since the 80 -84 MHz band was used for receive it would be very difficult to effect a change in its use whilst maintaining continuity. Similarly the 140 MHz region was used in the main, for base station receive links and the 150 MHz region for base station transmit links; again an inversion of direction of usage would be virtually impossible to accomplish without large scale disruption.

The bands finally adopted then were.

70.5 - 71.5 MHz Fire mobile, Base Transmit
80 - 84 MHz Fire mobile, Base Receive

143 - 144 MHz) Police mobile, and fixed links Base Receive
146 - 149 MHz)

152 - 153 MHz Police mobile, and Fixed Links Base Transmit
154 - 156 MHz

For either the police or fire situations, the small separations between the go and return channels will provide improved performance for the mobile aerial effect, compared to present operation with the wide spacing of 80 - 100 MHz. This factor should mitigate against possible degredations in performance due to the increased shadowing at the higher frequency and the effects of the more stringent filters needed to separate the bands.

9. Mobile Equipment Aspects

9.1 Introduction

The interaction diagram of chapter 3 (figure 3.2.1) showed the relationship of the design and performance of the mobile to the rest of the communication design. As with most of those other aspects, it was shown as being involved in a number of design feedback loops so that it was part of the chicken and egg situation of the uncertainty as to whether it should be determining the design of the interfacing elements, or have its design determined by them. This chapter will show that it is in fact largely a determining factor in the overall design, and also not very susceptible to influence by the system designer.

The reasons for this are probably part historical in that its performance, in its own right, has been optimised over a long period of time, and part the associated position it occupies in the overall economies of the system. Although of low unit cost compared to most other parts of the communications chain it dominates in numbers. Thus an emergency services system might have one control, four base-stations but perhaps 500 mobiles! The numbers involved attract the competitive aspects of manufacturers so that there are strong evolutionary factors involved which tend to ensure that mobiles have a near optimum cost/performance trade-off with the different manufactureres tending to occupy slightly difficult parts of the trade-off spectrum.

Although part of many loops so far as the system designer is concerned, to the user it is just part of the communication chain. But probably as far as he is concerned, the most important since it represents to him the last link in the chain and the one with the highest visibility to the largest number of users. Thus to the bulk of the users the performance of the communication system will be realised at the mobile, and it therefore has both technical performance on the one hand, and on the other ergonomic and visual appeal aspects to be considered.

Four main sub-divisions of the 'mobile' will therefore be considered. These are; the mobile aerial, the external interfaces of the mobile set, the internal characteristics of the set, and the physical aspects.

In all cases the mobile is taken to be vehicle mounted and to operate in a two way (ie transmit and receive) mode.

9.2 The Mobile Aerial

The influence of the mobile aerial on performance has been discussed in chapter 8 where the effects of operating at a non resonant frequency were described. It was stated there, that this meant that the aerial was mismatched to the feeder and that the receiver was likely to be mismatched also [the receiver mismatch will be covered in the next sub-section on interfaces (9.3)]. The combination of these was in effect to place a lossy filter in the path between the aerial and the mobile. Since the choice of frequency bands became favourable from this point of view, ie the go and return frequencies were close and therefore the aerial could be tuned to a good compromise frequency, the effect so far as the emergency services were concerned could be ignored.

The mobile aerial is almost invariably a quarter wavelength whip, since this gives, at least under 'laboratory test' conditions, very good results. The laboratory tests usually ensure that a test aerial is placed on a reasonably large ground plane so that that forms a reflector and the performance of a half wave dipole is simulated. This is only reproduced in a car by mounting the aerial on the roof at its centre. This ideal position is frequently not used however. The reasons are many, such as:- 'cutting a hole in the roof centre makes it difficult to sell the car', it is more important to have a flashing beacon or sign at the roof centre (this of course argues against the first point), 'they' won't lower it when going into restricted heights and it gets chopped off, 'they' forget its there when they take it into the automatic washer and it gets damaged, technicians do not like the bother of an extensive removal of the head lining of a car to make the installation. The consequences of these reasons is that it is often fitted on a wing or at the front of the roof. A report produced some time ago for the Directorate (Burberry and Smith 1974) showed that this made the installation non omni-directional and that there could be a loss of 10 dB or more in signal strength.

The fire service have problems due to the roofs of their fire appliance cabs often being non metallic, so that an artificial ground plane is required. There are also the difficulties of conflict with ladders and low clearances for appliance houses (Fire Stations).

In addition to the polar response aspects of the aerial, considerations have to be given to the fact that some locations will pickup more vehicle generated RF noise (ignition, motor etc) and also have longer feeder runs. Thus what at first sight is perhaps a mundane link in the chain is one which could well repay detailed study to improve its performance - electrical and otherwise. However the time and effort were not available for this - certainly not with the WARC time scales so any work will have to be deferred and put into the 'long term' category, and the conventional quarter wave whip will be assumed with a performance equal to that for current systems.

9.3 Interfaces

9.3.1 Audio

The mobile equipment has three main interfaces; that of the audio input and output, that to and from the aerial at RF, and the physical impact on the user. The latter will be dealt with in a separate topic in section 9.5 whilst the other two will be covered here starting with the first one.

Since the mobile equipment both transmits and receives (but only one or the other at any one time - see section 9.4) it will be necessary to consider at least two transducers the microphone and the loudspeaker.

The design of both these units is well outside the field of this thesis since it involves significant physical/acoustic aspects. But some comment can be made since they are the very first and last elements in the communications chain. The microphone is realised as a cord connected device, which can be unhooked from its resting place and brought near to the users mouth. This enables the best acoustic input signal to noise ratio to be achieved in what is, by any standards, a high ambient acoustic noise environment. The carbon microphone has survived for a surprisingly long period here, its poor performance in terms of distortion and noise being perhaps compensated for by its robustness. Its need for a bias current has not been a great disadvantage since power must be supplied to the microphone device to be controlled by the 'press to talk' switch which is microphone mounted for convenience of the users. Since it is a fundamental requirement of operation that the mobile be only enabled to transmit when this is positively desired, then the operator must use one hand for this. Thus holding a microphone is no extra penalty, as might at first be thought, compared to a boom mounted microphone.

The loudspeaker is usually dashboard mounted often as part of the control unit and, like the microphone performance, it should ideally be characterised in terms of its sound pressure levels. However this would require extensions into acoustic testing which is both expensive and more open to argument on its findings. This is particularly so when the performance of both the transducers have been 'tailored' by the manufacturers, by electrical and acoustical devices, to give what they consider to be the best sound.

In fact therefore the first and last links of the chain - the acoustical ones - are effectively not measured. All the measurements are taken at the electrical interface. These specify the impedances, levels, and frequency response in fair detail but leave the efficiency (electrical to acoustical power) and frequency response of the transducer largely to chance.

9.3.2 Radio Frequency

This interface is again a two way one. When transmitting, this is specified in terms of power, output impedance, and tolerance to mismatch. The power, for emergency services work is some 20 watts carrier which means upto 80 watts PEP. Associated with this are requirements for spurious outputs - mainly in the form of discrete frequencies caused by local oscillator or synthesiser/multipliers, or by more broad band noise. These aspects are well covered by the regulatory authorities specifications. Tolerance to extreme mismatch on the aerial is a practical requirement to cover either short or open circuits in the feeder or the absence of an aerial which could appear to the transmitter as anywhere between an open and a short circuit depending on the length of feeder.

Receiver input performance is an aspect also covered by regulatory authority requirements and, as mentioned in the introduction, competition between manufactureres is particularly keen here since this is seen as an important selling point. So that there is every chance of that regulatory performance being met and exceeded. However there are trade-offs to be made here between a number of factors such as sensitivity, adjacent channel rejection, and front end linearity (intermodulation response); and under ideal conditions this would be an area of activity for the system designer. Particularly when the consequencies of the frequency plan (chapter 12 and 13) show that the mobile receiver will have to tolerate both adjacent channels and, more pertinently, channels which will raise the receiver mute due to 3rd order intermodulation products being generated in the front end. Thus the detailed frequency assignment plan should have an impact on the design or specification of the mobile receiver.

9.4 Internals

This section is intended to cover all the aspects of the mobile equipment between the input and output terminals.

The author has discussed the design of mobile radio receivers elsewhere (Fudge 1984b) and again the inter-manufacturer pressures will tend to ensure good performance and design. In fact any short-fall in the field would be brought out - even if not expressly given in the equipment specification - during evaluation of competing models.

There are however some aspects which could be deemed special to the Directorate's needs and therefore worthy of attention. That of the trade-off between input parameters has already been mentioned. But, it will be recalled from chapter 7 on modulation, that the ALC and mute actions will be important for the quasi synchronous mode of operation chosen. These parameters are not usually specified in any way except to call for a user-variable, or technician pre-seltable, mute level. The quasi-synchronous operating parameters of section

7.8.6 were derived for one mobile equipment which had not been optimised in this respect. The mute and ALC time constants were those thought best for opration in the normal Rayleigh fading environment. Ideally then, these receiver parameters should have been investigated to determine their optimum values, these values specified for future receivers, and the overall parameter determination of section 7.8 repeated for these values.

Time and lack of resources did not however permit this ideal course and it remains an area to be investigated.

It is normal to specify a mobile receivers performance in terms of its 'sensitivity' this is taken as the level of input carrier to achieve a specific SINAD value - often 12 dB. Whilst giving a measure of the sets low signal performance it does not necessarily equate to a particular degree of performance at the more normal higher input levels. For the case under consideration therefore the opportunity was taken to specifying minimum SINAD performance at three input levels. These are shown on figure 9.4.1 as ABC.

This figure is a reproduction of the AM portion of figure 7.42. Point C tests for the linearity of the demodulator and audio amplifiers and is additionally measured under conditions of high modulation depth (80%). It demonstrates the best performance which can be expected from the system. Point B is intended to lie on the straight line section of the graph where 1db of variation in carrier input level causes 1db of audio signal to noise ratio variation. It should be dependant on the noise figure of the receiver only since it is clear of the non linear performance associated with the region around C and the detector performance around A. Thus a calculated value can be specified to force a desired noise figure.The point A is given in region of the normal sensitivity performance figure, but in this case it was sensible to riase it somewhat above the unusable) 12 dB SINAD figure but still leave it on the region of increased slope caused by envelope detection since it would effect quasi-synchronous performance.

Another area touched on in the modulation chapter was that of signal processing, particularly that of compression and expansion (companding) this again would seem to be a topic which would yield good returns for the AM systems under consideration where it does not seem to be used at all.

9.5 Physical Aspects

The author's already cited publication on receivers (Fudge 1984b) points out that the mobile eqiupment operates in what must be considered to be a very hostile environment.

The high acoustic noise levels have already been mentioned, and so has the electrical noise generated by nearby devices such as ignition, electric motors and invertors, and possibly tyre and brake static; together with external factors of poor signal level, which is fluctuating rapidly, and external noise from other vehicles and industrial premises. In addition it is expected to operate over a very wide range of temperature (-10°C to +60°C) and to do this either virtually continuously or immediately after switch on. It is subject to vehicle motion forces and vibration, and will be refitted into several vehicles whose lifetimes are shorter than that of the radio.

It will be mounted in a way which is most convenient, not for the performance of the radio, but for other aspects such as space available etc. Since it is larger than the space allowed for in the design of the vehicle for 'in car entertainment' devices, then it is difficult to mount in the passenger compartment of the vehicle (if put under the drivers seat for instance it is liable to be kicked to pieces by people under arrest). Therefore a remote operated control unit is mounted in the entertainment space or close to the driver and this contains the channel selector, loudspeaker and microphone - or sometimes a telephone type handset with microphone and earpiece for moderately private reception. The main set might be mounted in the boot where it may not experience the extremes of temperature (since it will not be subject to the greenhouse effect due to the windows) but it will have to compete for space with tools, signs, traffic cones etc).

All these physical factors are of importance, not directly to the system designer but because they do have an influence on customer reaction. There is then, in addition the aspect of presentation to the user, ergonomics, and not least the question of maintenance.

9.6 Conclusions

As stated in the introduction to this chapter the system designer will in general have little impact on the design of the mobile due to the strong optimisation of competition. However it has been shown that the mobile receiver design will interact strongly with other parts of the system design and, that for the use envisaged here, there are elements which would ideally require specific optimisation - specifically the ALC and mute, and the signal processing aspects.

Aside from these the system designer can influence the evolution and competition by calling for sensibly stringent performance specifications. Notable in this respect are defining receiver performance for several RF input levels (at maximum sensitivity, at high input levels, and an intermediate one where the performance can be calculated). This performance covered also the audio frequency range to an extended requirement, and the measurement of distortion under realistic conditions (SINAD). These aspects could be considered to be interactive in that they would admit of relaxations elsewhere for the same overall performance. However, it was considered better not to aim to take advantage of this for the system design but pass it to the customer in the form of improved performance.

10. Base Station Design

10.1 Functions

10.1.1 A Simple Base Station

The base-station at first sight appears to be just another element of the communication chain and, due perhaps to its physical remoteness and immobile appearance (as opposed to the control room which has high visibility and activity), one of low importance. From the system design point of view this is totally untrue and the fact that it occupies a central position on the interaction diagram is some way indicative of its central role.

In its simplest terms it serves to take signals from a control, radiate them to a mobile, receive the mobile's return signals and relay them back to control. To obtain best coverage to reach mobiles it is located on as high a spot as possible, and then has a high tower to carry the aerials. The fact that it is a high spot implies remoteness from control and that it is in general isolated. Therefore the links to and from control are best implemented on radio carriers. Land lines would be expensive to instal, subject to breaks in operation, difficult to repair, in general more costly to implement, and perhaps less private. Thus the simplified diagram of a base-station, which can also be referred to as a hill top site, is shown in figure 10.1.1.

This shows the simplified situation for one circuit (a go and return channel on both main and link). It shows the two main parts of the hill top site:- the tower with its aerials, and the equipment room. It does not show, even for this simple case, the built-in reliability achieved by having duplicates for each active piece of equipment which can be brought into operation, either locally or from control, in the event of a failure of the working module. Such an operation is termed a 'bay change'.

10.1.2 A More Complex Base-Station

Figure 10.1.2 represents the situation for a more complex, but far from a highly complex, one. It shows the base-station (B) connected by radio links to police (P) and Fire (F) control rooms depicted by triangles. Two police controls are shown on the assumption that this is a site which has advantageous cover for two counties, it thus becomes a chaining site as discussed in chapters 12 and 13. Additionally a link to another base-station is shown which makes the one under consideration a repeater site also. The numbers against each line represent the number of channels (each two way) carried on that path.

The equivalent equipment diagram is shown in figure 10.1.3. This is basically just a multiplication of the single channel case with the frequency bands of operation shown. The repeated link poses problems, the solution to one of them being shown on the diagram. The extra equipment involved in this repeat path will cause the modulation at the onward linked site to be out of step with the others. This is unacceptable for the reasons given in chapter 7, so to rectify this all the other transmissions will have to have a compensating unit which simulates the characteristics of the extra link transmitters, receiver and associated equipment. This can be considered to be equivalent to a delay and a frequency dependent phase shift. Units to supply this are shown in the diagram. The equipment to equalise the differences in other, non repeated, quasi-synchronous channels is best located at Control where the adjustments are best made.

The other problem with repeated links is that the next site, shown as B' on the diagram, will in effect have this link in reversed frequency bands, that is it will have a link receiver in the 150 MHz band and a link transmitter in the 140 MHz band. This will cause difficulties for frequency assignment and is discussed in chapters 12 and 13.

Another factor shown by the diagram is that of access by another control - probably in an adjacent county. This will interact, so far as this chapter is concerned, with the main user of the site through the general increase in equipment and particularly the impact of intermodulation products which this will engender. It will also have a significant impact on the frequency planning - again part of the chapter 12/13 discussion.

It must again be emphasised that the situation shown here is only more complex in the sense that it is handling more services. Many aspects are not shown either because they have yet to be discussed or because they are not relevant in these initial passes of the interaction diagram. They include, filters redundant equipment, supervisory and monitoring, standby power supplies, high stability oscillators etc.

The introduction would not be complete without a mention of the operating conditions - whilst onerous they are not as severe as for the mobile. The equipment must however operate unattended for many days or frequently weeks at a time, this is reflected particularly in the need for very high stbility frequency sources to achieve the nominal 1 Hz difference between remote quasi-synchronous carriers. Although specified to operate over a wide range of humidities the heat dissipation of the equpment room is usually such as to minimise the total excursion. This factor also, in practice, reduces the extremes of the temperature specification of $-10^{o}C$ to $+50^{o}C$. The remote location also gives rise to the wide range of electrical supply voltage which must be tolerated together with the ability to withstand surges, whether caused by lightning or even a complete failure - change to standby - and recovery.

10.2 Aerials

10.2.1 Radiated Powers

The types of aerials used and their dispositions are closely connected with the arrangement of the transmitters, and both effect the radiated power. This would be true for the single aerial per transmitter case shown in figure 10.1.3, and is even more the case for the considerations to other arrangements which will be given here. It is very much another chicken and egg argument as to where the starting point is. In fact the interaction diagram shows that once again an interactive design loop is involved, and in fact this was traversed many times in the course of this study with various configurations before a compliant solution started to emerge.

To follow the actual sequence of events would be repetitious and wearying so an attempt will be made to present the arguments in as much a serial form as possible, but it must be recognised that this will not be able to cope with all the configurations employed, and may well leave gaps in explanations which hopefully will be covered subsequently.

The arguments which follow are those which apply to the main transmitters and receivers not those of the links which are assumed at this stage to remain as they are; that is, general use of six element yagi aerials possibly on orthogonal polarisation for go and return.

The starting point for the determination is that of the required radiated power, as shown in the interactions diagram. In the old bands transmitters with nominal powers of 50W (60-70W actual) are in general use with a few 250W cases. These feed half wavelength dipoles. For the new bands the surveys of chapter 4 provide apparently a firm and rigorous basis on which to determine the figure which should be used. Arguments in that chapter and appendix A show that 17 dBW (50W) radiated from a dipole should be acceptable. Although this matches well the figure used before, there is an element of artificality in it due to the assumptions which have been made about 'typical distances' 'typical county' etc. It is necessary therefore to further justify or modify this figure.

This can be done by building on experience of performance in the existing bands. It has been stated before that the system has evolved over a number of years to be compatible with the users desires, as modified by the size of his purse. Present coverage is therefore at least adequate - with a few exceptions where extra base-stations are the real solution. If the same powers were to be radiated in the new bands, then the arguments of the previous propagation chapters show that performance should be largely similar with the slight reduction in performance due to shadowing/foliage at the higher frequencies being compensated for (at least in the outgoing direction) by improved mobile aerial performance. But the 80/100 MHz system has poorly designed aerial systems (usually a simple dipole on the side of the tower) and moderately lossy feeders. If the new system can demonstrate the overall cost benefits of lower loss feeder and better aerials then such benefits can either result in:- improved performance (higher erp and/or aerial gains), or reduced transmitter power or receiver sensitivities. There are some strong arguments for keeping transmitter powers low from the standpoint of intermodulation generation (to be dealt with in section 10.4 and chapter 11) but it was decided that overall the improved performance should be treated as a 'reserve' to accommodate any unforeseen factors which could themselves degrade performance.

Therefore a standard of 17 dBW (50W) of power radiated from an omnidirectional equivalent dipole was taken as the objective.

10.2.2 Polar Diagrams

The existing 80/100 MHz system in general uses a quarter wavelength unipole (a ground plane) at the top of the tower as a receiving aerial, and dipoles mounted some specific distance from the tower side for the transmitters. The ground plane is connected to an aerial distribution amplifier which is wide band and feeds in turn the receivers. The ground plane is generally assumed to have a constant azimuthal response - circular polar diagram, and measurements of true responses from actual towers have largely confirmed this.

Theory, and manufacturers catalogues, purport to show that the circular polar diagram of an isolated half wave dipole is little affected if spaced a particular fraction of a wavelength from a conducting rod -

the assumed supporting scaffold pole. It is this spacing which has been used for tower mounted dipoles in the somewhat naive believe that circular coverage will result. The naivety is however tempered by ensuring that the dipole is mounted on the face of the tower which corresponds to the direction of desired best signal, or at least the dipole has a 'clear view' in that direction. The same measurements as were carried out on the ground plane, show that this is a wise precaution since the response can be 20 dB worse in other directions - usually that opposite to the dipole mounting. This is hardly surprising in view of the large size of the tower cross-section in terms of wavelengths. The tower will be acting as a reflector or number of parasitic elements, the overall response being as ill-defined as is the tower structure and the exact relation of the dipole to it (eg tubular structure or angle section, position, direction and size of cross braces, length of mounting boom, position in relation to cross braces, etc).

Omni directional cover (circular polar diagram) was therefore decided on for the following reasons:-

- It ensures that there is no degradation in cover compared to the existing system (there will be some enhancement in the direction of the old system's notches).

- It makes the task of the frequency planner easier when determining reuse distance.

- To cater for all individual responses would be too demanding on planning resources.

- There will not be time to optimise individual patterns during installation.

- It admits of common aerial working.

If only a few channels were to be radiated from a site (say 2 or 3) then the expense of a common omnidirectional aerial might not be bearable and dipoles would probably be used.

10.2.3 Constraints on the Realisation

The work on aerials for the WARC programme formed an extension of that already in-train as part of general system improvements and which were in progress under the author's control before the start of this thesis. In addition the direct control of the work on aerials (and directly on transmitters/receivers) passed from the author during the period of the thesis. In consequence the information which will be presented here will not contain the details of the work, with its large number of theoretical and practical polar responses, but rather an overview of the work and its impact on base-station design and other system design factors.

Any nominally omnidirectional aerial mounted on the top of the tower (eg dipoles, end fed dipoles, ground planes, colinears, turnstiles etc), would in principle meet the omnidirectional requirement. It would then have to meet the total transmitter power requirements for both average and peak conditions and for amplitude modulation this could be onerous if there were only one radiating element as in the case of a dipole or ground plane (10 channels at 50W requires a 500W average capability and 2kW peak power capability). Additionally only the transmit or receive aerial could occupy the generally sought after position (if the tower was not owned by the local authority but we were a lodger then this position would almost certainly be in use). Duplexing to admit transmit and receive on the same structure was not possible with practical filters. On this basis if any aerial occupied the tower top it would be sensible to make this the receive one. But this position was that which was most prone to lightning strikes - which could take the aerial serving all the transmitters out of action to say nothing of the receiver distribution amplifier. It is also the position where the incidence of precipitation static was highest. Aerials mounted further down the tower were in effect shielded by the upper parts of the tower discharging any damaging fields. It was therefore growing practice to not use the tower top for receive, if at all, at the start of the WARC planning period and it was decided that this should be the aim of the new aerial system.

10.2.4 Aerial Configurations

The aerial system is then to be located below the top of the tower and it has already been shown that a single dipole will not meet the omnidirectional coverage if located at this position. Therefore a more complex arrangement is needed and it is fairly obvious that a number of aerial elements will be needed - distributed around the tower at the same level.

This is shown in figure 10.2.1 which depicts the plan view of an array of such aerial elements. Each element is connected by a feeder to a distribution network or combiner shown as C. Tests were conducted on such an arrangement where the elements were each a dipole, and eight elements were used. The combiner consisted of a cascade of coupling harnesses as described in appendix H. The results were disappointing and the reasons ascribed to this are discussed.

The simplified view shown in figure 10.2.1 suggests that each element can be considered as a radiator, if there are enough of these and they are all fed in phase then each can be considered as a Huygens radiator on an advancing circular wavefront central on the axis. Since the individual dipole radiation patterns are circular in the plane shown (the horizontal) then such arguments hold. Only in directions other than the horizontal need the true dipole pattern be considered. The first question to be settled is whether there are enough of them. This depends how far apart they are. When the spacings become significant then the individual wavefronts must be considered together with their path length differences in the direction to be evaluated, as in normal aerial array theory. On this basis suitable spacings can be determined which do not markedly detract from true circular coverage and this can be confirmed by computer generated plots. But this still leaves two other effects, the first of which is the coupling between array elements, particularly the adjacent ones, which will be apparent as a mismatch at the element to feeder interface and, due to the symmetry will also appear at the combiner to main feeder interface. The second is that of the influence of the tower, where the real situation is shown in figure 10.2.2. It will be seen that there is now four-fold symmetry due to the tower not the eight-fold envisaged. The tower also effects the individual radiation patterns (and impedance matches). Even with a sheet metal tower the overall response would be difficult to predict, with a lattice tower far more so.

The first step in overcoming these factors was to use aerial elements which effectively isolated the active element from the tower. For this, Yagi arrays were used for each of the original elements. A six aerial array was tried but again results were a little disappointing and problems were envisaged with the wind loading of the array. The disappointing performance was ascribed to insufficient control being available over the location of the phase centres of the aerials to overcome the interference effects in the areas where the beam widths overlapped; and there was still significant coupling between the aerials.

Since the tower imposes a four-fold symmetry on the final pattern it was concluded that this should determine the array shape and four banks of dipoles backed by a reflector were considered, but these gave way to, what was physically and electrically similar, four skeleton slot panels. Each panel consisted of a reflector grid of vertical rods with an arrangement of two or three spaced dipoles with their tips folded to connect, thus forming a rectangle and termed a skeleton slot. These performed very well in trials at 80 and 100 MHz, and could have been made broad band enough to cover both those bands if desired. Due perhaps to their relative position the coupling between them was very low (40 dB) and this, coupled with their inherent lightly controlled quadrant coverage, gave the good results.

This then was the type of aerial structure envisaged for WARC system planning, a ring of four skeleton slots for receiving, and another ring (or two rings - see section 10.4) for transmitting. Both these would cover the 140/150 MHz requirements. The low numbers of channels for the fire service requirement at 70/80 MHz being achieved by single aerials - dipoles with the consequential coverage pattern. For the sites with only a few 140/150 MHz channels consideration would be given to the use of a single dipole per channel.

For the major sites therefore the transmitter could be presented with upto four or eight feeders and the receiver distribution amplifier with four. For the lesser sites each channel would have its own feeder.

In fact during the experiments and trials with the 150 MHz slots it was found that the old trouble of phase centres reappeared. The tower cross-section were just too large to admit of the close spacing desired for the wavelength, and at the time of writing attempts are being made to overcome this. It is after all part of a continual development.

10.3 The Combiner

10.3.1 General

If each aerial is connected to only one device such as a receiver or a transmitter then that connection is obviously a simple one. If however one or more aerials has to be connected to one or more devices then consideration has to be given to the splitter or combiner required.

In the case of the receivers, as already discussed, one aerial usually feeds a number of receivers and this is accomplished by a receiver distribution amplifier. This is of course an active device consisting of a low noise, moderately wide band, input stage which if it has enough gain will feed a resistive splitting network. For extensive splitting, with a large number of outputs, either an extra stage of gain is used or the first distribution amplifier is followed by others. The aim is to meet the normal receiver's linearly requirements and to make little degradation to its noise figure. The device is of generally proven, although still developing, technology.

For the transmit side the story is more complex.

10.3.2 Transmitter Combiners - Passive Simple

The simplest case is that of combining two transmitters onto a single aerial (or single feeder of an aerial system) this is shown in figure 10.3.1 with C again representing the combiner. At first sight this could be the coupling harness of appendix H but it is shown there that the isolation between the input ports is not high and this will lead to intermodulation product generation by the mechanism to be described in chapter 11. It is therefore unacceptable here.

An alternative is to make C take the form shown in figure 10.3.2. Here filters are used and, as shown, they are band pass, each tuned to the frequency of its transmitter. By suitably choosing the lengths of the cables a and b, the Tx1 branch will appear at J to be an open circuit to signals from Tx2 so all its energy is coupled into the aerial without loss. Similarly for Tx1's signals. This then is the conventional filter multiplexer (in this case a diplexer). It requires however that the frequency of Tx2 be reasonably well into the stop band of filter F1 and vice versa for the open circuit condition to apply. Using the highest Q filters available (in fact cavities made by silver plating suitably cut-down beer barrels) the spacing between nearest channels could not be less than some 300 kHz at 150 MHz. This figure applied even if band reject filters (notch filters) were used, tuned to the complementary frequency. Their use would then severely constrain the detailed frequency plan (see chapters 12 and 13) and in any case the filters would have to be tuned on site and maintained in the correct state. More channels could be combined in a virtually lossless way using this technique but all spacings would have to be 300 kHz. Therefore, although they could meet the coupling requirement for intermodulation generation, they cannot be considered for general planning purposes.

Another alternative for the combiner is the use of so called hybrids or, perhaps more properly, couplers. The diagrammatic representation is shown in figure 10.3.3.. The coupler may take several forms of which the stripline, Wilkinson, and ratrace are some. The resistive load, shown separately, is often incorporated in the device when it is intended to be a three-port and is therefore left off the diagrammatic representation. The device works well, in that they have inherently a bandwidth wide enough for the purposes being considered here, and can provide good isolation between input ports (30 dB - or more with tuning). The price to be paid for these properties, which allow the transmitters to be on adjacent channels if desired, is that of attenuation. Half of each transmitters signal is lost into the resistive load if the coupler is symmetrical. They can, with occasional advantage, be made assymetric so that one source is little attenuated and the other more so.

The loss of signal, although only 3 dB, is not to be taken lightly since it has been expensively generated in the first place. In fact to achieve the desired 17 dBW radiated from a dipole then each of the two transmitters would have to be some 130W capability to take account of the 3 dB loss of the coupler and the nominal 1 dB of the feeder. Such transmitters are not generally available.

A solution to this for the single aerial case might be to have the feeder supply two dipoles in a suitable vertical arrangement, thus having a stacked array with an overall gain of 3 dB.

10.3.3 Transmitter Combiners - Passive Complex

If it is desired to couple a number of transmitters to a number of aerial feeders such that each feeder receives an equal contribution from each transmitter then a cascade of coupling devices is needed as shown in figure 10.3.4. Here the devices marked H are either hybrids as shown in the lower part of the diagram, or the coupling harness of appendix H if their inherent mismatches can be tolerated. In either case the loss from one transmitter to one aerial output port is 12 dB, although the net effect, if each aerial element gives quadrant cover, is a loss of 6 dB compared to one transmitter supplying a well located dipole.

An alternative to the arrangement shown in figure 10.3.4 is that of figure 10.3.5 which is the so called Butler matrix (Butler and Lowe 1961). This relies on a particular type of hybrid (180° type), and control over the lengths of the interconnections of the hybrids. It does however provide a device which is lossless in itself. Each transmitter experiences a loss of 6 dB to each aerial port but this is recovered by the effective gain of each aerial element. The device is inherently of broad band performance (so far as this work is concerned) since the only frequency concious parts are the interconnecting, and internal hybrid, line lengths. Its major disadvantage is that it is difficult to incorporate more than four channels. It also has a small impelementation loss of some 0.5 dB per traversed hybrid, giving an overall loss of 1 dB.

Combinations of the simple techniques could be used such as that shown in figure 10.3.6 where A and B must be more than 300 kHz apart as must C and D but there are no restrictions on the assignments between the two groups.

Thus passive combiners can be used either, where the losses or frequency restrictions can be tolerated - which usually would mean at a site having only a very few channels, or where there are not more than four channels by using a Butler matrix.

10.3.4 Transmitter Combiners - Active, Initial.

Many sites will need to cater for significantly more 150 MHz channels than those admissable by the techniques of passive combining. Such large numbers of channels however make it certain that the aerial array can be extensive.

The simplest form of active combiner is to take say the hybrid cascade of figure 10.3.4 and add amplifiers to it. These can be located in several places. If located at the outputs to the aerials then four of them will be needed, each capable of handing the four channels of this case, without the production of significant intermodulation products - this is an onerous task.

They could be located at the inputs to the networks where there would be the enormous advantage of having to deal with only one channel. But the power rating of all the subsequent components would have to be increased significantly, and the power generated is being dissipated. The third location for the amplifiers is within the network, either as a single device at the node between the combining section and the splitter, or as a pair after the first splitter. The latter has the advantage of not putting all the eggs into one basket.

Experiments were conducted in the periods before the start WARC programme with the first two options. It was shown that semiconductor devices could not meet the performance requirements (on power grounds let alone general linearity or intermodulation aspects). So that thermionic valves were necessary with all their attendant high voltages and power requirements. The gain per stage was some 12-15 dB. This was barely adequate to cope with four channels of combining, but another stage would have been something of an over-kill from the gain point of view, and introduced matching problems which would have resulted in narrower bandwidth of operation than was desirable. Furthermore it relied on the delivery of relatively high power from each of the transmitters - effectively that of a full power transmitter. On this basis incoporating active devices would be expensive.

10.3.5 Transmitters Combiners - the Multichannel Transmitter

A derivative of the third method of producing an active combiner was developed into a device called the multichannel transmitter.

The basis of the technique was to combine the separate channels at low level and at a low frequency (in fact some 30 MHz). A judicious degree of amplification could then be obtained with solid state devices before all the channels were up-converted to their final frequencies at VHF. The up converter was followed by two lots of thermionic amplification, firstly an intermediate power amplifier (IPA), and secondly the high power amplifier (HPA) in the form of a carefully chosen valve. The form of the transmitter is shown in figure 10.3.7.

Each of the two outputs was intended to feed one tier of an array of skeleton slot panel aerials. The two tiers would have a gain of some 5 dB and allowing for feeder and harness losses each amplifier output needed to have only 20W per channel capability. Even so, a 1 kW capability was needed for the HPA to achieve the desired level of 3rd order intermodulation products 55 dB below carrier, fith orders 65 dB below, and higher orders at least 70 dB below carrier.

The multichannel transmitter was capable of handling upto 12 channels with any desired spacing - even adjacent channels. Moreover the channels could be of any modulation standard, such as AM, FM, SSB, and this could be easily changed by a change of just one small module - the modulator card. Reliability was obtained by having all the devices carrying more than one channel, in effectively working stand-by. With such a device it was also economic to use a redundant pair of very high stability frequency sources so that the quasi-synchronous offsets could be easily set and maintained.

The experimental models were taken by industry and developed into pre-production evaluation models which satisfied the requirements very well. This was for 100 MHz working however and the translation to 150 MHz devices was never completed due to a large number of causes only a few of which were, relatively minor, technical problems. Therefore this device is scheduled to be used as a temporary facility at each hill top site carrying all the existing 100 MHz services on a temporary aerial structure whilst all the permanent 100 MHz carriers are replaced with 150 MHz ones during the actual changeover. When one site has been finished it will move on to the next where its inherent flexibility will enable it to be quickly brought on line.

For the main 150 MHz combining improved single channel transmitters and easing of transmit-receive frequency spacings allows general use of devices such as that of figure 10.3.5 (the Butler matrix).

10.4 Transmitter Design

10.4.1 General

There are a number of aspects of transmitter performance which are relevant to their operation in any circumstances such as; discrete spurious outputs, output noise, over modulation, frequency stability, etc. Where the performance is not significantly different to such general operation or is covered by industry wide blanket specification, for example the MPT series of specifications, then no specific comments will be made here. Rather those aspects which are particularly relevant to the WARC exercise, are novel, or of possible relevance to other mobile radio users operating in a similarly congested mode, will be highlighted here.

For this exercise the output from a base-station was set at 17 dBW (50W) from a half wave dipole or its equivalent. The last section commented on the fact that this equated to 20W output from each half of the multichannel transmitter. If a single channel per transmitter is considered - either to feed a dedicated aerial or via a combiner such as the Butler maxtrix (driving a four element skeleton slot array) - then the total losses between the transmitter and aerial were set as a system design parameter as being 3 dB maximum. Therefore the output power from a transmitter needs to be 20 dBW or 100 watts.

The frequency stability requirements were mentioned above and the value to be used here is that derived in chapter 7 on modulation for quasi synchronous operation. That requirement was for a nominal offset between base-stations of 1 Hz. This then becomes the requirement for any transmitter which might be called on to operate in this mode, if not operating quasi synchronously then ordinary MPT specifications are adequate. To achieve this very high stability, for periods of some 3 months considered necessary between maintenance adjustments, a dedicated frequency source is needed rather than that which would otherwise be found in transmitters. It becomes economic then to provision for tranmitters which will either 'free run' with their normal frequency source (for use on non QS channels) but with the ability to take a (high stability) reference as input and derive the final frequency from that. In any case, moves in that direction were under way at the start of the WARC programme for both fixed and mobile equipment so that only one reference frequency (crystal) would be necessary for each range of equipment, the final result being derived by synthesis. The ability to reference to an external source therefore represents only a small change in any design.

Since every base-station will most probably have at least one QS channel in operation then it becomes sensible to think in terms of using the very high stability reference source (provisioned as a redundant pair) as inputs to all the non QS transmitters. These will be capable of taking such an input due to the bulk buying needs to make all equipments identical. There is a possible danger here however in that the tranmissions may be phase coherent and, where intermoduation products are concerned, then two or more products derived from the same difference frequency (even if from entirely different transmitter frequencies) will be identical in frequency and should therefore be summed on a voltage basis. This is dealt with in chapter 11 on intermodulation.

10.4.2 Output Stages

The decision to adopt amplitude modulation took into account the implications for the transmitters and in particular for the output stages of them. In the simplest form this would have to operate class A to achieve the required linearity of performance (as far as the modulating signal is concerned). It was confirmed that solid state output stage with the required performance could be designed and manufactured. [Other (world-wide) users of AM are the aeronautical services and there is a continous demand for such equipment]. From the system point of view the use of such inherently linear output stages was expected to be beneficial from the generation of equipment based intermodulation products (see section 10.5), compared to those which operated class C as in general use for FM. [In fact the AM transmitters use class AB but this seems to make little difference].

Alternative forms of output stages were proposed which used class C operation for AM. The RF output of such a device was sampled in the transmitter by a directional coupler, and a demodulated version of this fed back to the input of the modulator. Thus a negative feedback loop was established and high modulation linearity claimed. The simplest forms of this had in effect only envelope feedback more complex ones had full orthogonal feedback and were advocated for SSB.

Claims were made (Petrovic 1983) that such transmitters reduced intermodulation product levels for the case where the unwanted RF signal came down the aerial feeder as will be discussed in section 10.5. Such claims are however misleading, in that the intermodulation reduction (compared to convetional non feedback transmitters), only applies for the case where the difference between the output and impinging RF frequencies is within the loop bandwidth of the feedback device. This means effectively within one, or at best a very few, channels. This is a very restricted field for intermodulation reduction since most channel separations will be greater than this.

10.5 Intermodulation Aspects

10.5.1 Mechanisms of Generation

At the start of the planning for the WARC programme it was recognised that there were two classes of mechanism within the base-station which could generate intermodulation products. These were firstly the inherent and relatively gross non-linearities of the active devices of transmitters and receivers, and secondly the passive devices such as filters, couplers, feeders, aerials, tower etc. Attention was concentrated initially on the high level effects of the active devices and this will be described in this section. But the passive effects were later shown to pose a much bigger overall problem due to the difficulty of overcoming or coping with them. So those effects, and the general topic of intermodulation will be covered in chapter 11.

Figure 10.5.1 shows the basic mechanisms by which the unwanted signals may be coupled into one of the active, and hence non-linear, stages. They fall into three main categories; direct coupling between the equipment mounted in the racks possibly by case leakage - designated C, coupling of the signals in the feeders - designated F, and coupling between the aerials. In the first of these the couplings can take place between many possible sources and sinks and at any of the frequencies within the equipments. For the latter two however the critical elements are either the coupling between the transmitter output stages to produce the intermodulation product (intermod) and then the coupling of this to the receiver where it might be an interference, or the direct coupling of the transmitter frequencies into the receiver input stage and the production of an intermod there.

The first two mechanism C and F are in principle under the control of the design and specification of the equipment. For the systems designer it is enough to know that sufficient isolation can be achieved between the devices for either mechanism for it to be a sensible requirement to be included in the specification. This was in fact done in these cases. The coupling between aerial feeders was suspect in the case of braided cables. This, combined with the fact that such cables were a potential source of passive intermods (the contacts between braiding strands are thought to be the cause) and that lower total attenuation was required than could be sensibly produced by braided feeder, lead to the adoption of solid outer feeder cables, and these have inherently very good isolation. Tests were also carried out to show that even if transmitters and receivers occupied adjacent positions in the equipment rack then suitable screening could be introduced to ensure that the intermodulation performance was acceptable. Thus the proof that the objective performance could be achieved was deemed sufficient to pass these aspects to the equipment manufacturers by way of suitable performance specifications.

10.5.2 Aerial Coupling of Active Devices - Transmitters

The first mechanism of aerial coupling induced intermods takes place in the output stages of the transmitters. Using the diagram of figure 10.5.1 it will be seen that the coupling between transmit aerials (or it could be the coupling between input ports of a combiner is A1. Therefore transmitter 1's signal will be reduced by this amount plus the feeder losses by the time it reaches transmitter 2's output stage. After conversion to an intermodulation product, the level reaching the receiver will be further reduced by the feeder losses and coupling factor A3. Knowing these coupling figures and the intermodulation performance of the transmitter the levels can be predicted.

For planning purposes it was decided that in the case of single aerial operation the figure for A1 should be 20 dB minimum and A2 or A3 should be not less than 30 dB. These then would accommodate the combining devices which would have better performance.

Since no equipment of the final form of the transmitter was available for measurement during the planning stages, then it was decided to convert an existing (valve) transmitter to operation at 150 MHz and measure its performance. The results are shown in figure 10.5.2. This shows the intermodulation level as a function of level of unwanted signal. In fact the level is presented in a form which corresponds to that traditionally used for 3rd order intermodulation performance. For that the transmitter's output stages are treated as a mixer with its output forming the local oscillator. The impinging unwanted signal from transmitter two is equated to the normal mixer input signal and the output is the mixed signal. This is represented diagrammatically in figure 10.5.3. This has the advantage of giving a single figure for intermod performance for 3rd order generation.

This perhaps requires some explanation. If the non linearity which causes the intermodulation is represented by a polynomial transfer characteristic (as will be discussed in detail in chapter 11) then the products of main interest to this exercise are those which derive from the odd order terms. These are the ones which produce products close in frequency to the incident signals. The amplitude of such a product is an n^{th} order function of the input amplitudes. Specifically if the two input amplitudes are r and s then for the products of interest the product amplitude will be equal to

$$\frac{r^p . s^q}{k_n}$$

where $p + q = n$ and, $|p-q| = 1$ and k_n is a proportinality constant

For third order therefore, $p + q = 3$ and either $p = 1$, $q = 2$ or

$p = 2$, $q = 1$

If r and s are equal then for the third order case the output level will be proportional to the cube of the input level, or if expressed in dB terms the output variation (in dB) will be three times the input variation (in dB). This is of course the elementary result. The situation considered here, however, is that one of the inputs (that of the signal passing through the transmitter) is constant. So again considering 3rd order, p can be considered constant in which case one of the intermodulation products will be dependent on q^2 and the other on q. Therefore one will have a dB variation of twice the input and the other will have the same variation. Since the latter one will be, in our case, the one with highest level, then attention is concentrated on it as the worst possible case, and the proportionality constant k_3 represents the 'conversion loss'. On the graph shown this will appear as a constant value for the third order product.

For the fifth order product, the highest level product will be a square term which if plotted on the same basis as for the third order case, will have a dB slope of 1. The results obtained in figure 10.5.2 bear out these assumptions and thus confirm the viability of applying a polynomial model to the production of intermods. Not withstanding the foregoing paragraphs however, it is in most cases easier to deal in absolute values rather than the conversion loss concept and the same results as plotted in figure 10.5.2, are shown in figure 10.5.4 in a form which is of more use for orders higher than third.

The results obtained here were used for planning purposes, as discussed in section 10.6. Although not necessarily representative of the performance of the final equipment to be purchased, it showed what could be achieved and it was therefore used to set performance requirements by way of specification. If manufacturers found that their transmitters were inherently worse than this (due perhaps to using solid state devices and the higher output power requirement) then this could be accommodated by including a final attenuator (whilst still achieving the required output power) or more probably the inclusion of an isolator.

10.5.3 Aerial Coupling of Active Devices - Receivers

The second mechanism of aerial coupling induced intermods, is that where the intermodulation takes place in the receiver itself. Here the two (or more) transmitter signals are coupled to the receiver aerial with couplings A2 and A3 (see figure 10.5.1). The front end of the receiver is not perfectly linear and therefore the intermods will be generated, and their levels can be predicted for various aerial couplings if curves of intermod levels were to be produced of the form of figure 10.5.4.

There were several problems in doing this however. Firstly no receivers, or receiver distribution amplifiers, of the type likely to be used at 140 MHz were available and it was judged to be not meaningful to extend any results obtained on the currently used 80 MHz ones to this frequency. The reason for this was two fold; the active devices at the front end were likely to be different, and the impact of the transmitter frequencies would be heavily dependent on the degree of filtering used on the input stages (furthermore there was an uncertainty as to the likely receiver frequencies, somewhere between 143 and 149 MHz, and similarily with the transmit frequencies, between 153 and 156 MHz). Secondly the level of the intermodulation product could not be measured directly - only its impact on a modulated RF signal to which the receiver was tuned - ie manifest as a degradation in performance such as SINAD.

A third factor was that it was recognised that the system would require filters external to the first active device, whether receiver or receiver distribution amplifier, for requirements other than those of intermods, and these filter characteristics would heavily affect the levels of the transmitter signals reaching the first active stage.

It was therefore decided that all these factors should be covered by specifying the intermodulation performance to be of a particular level with these filters taken into acocunt. This is discussed further in the next section.

10.6 Transmit-Receive Compatibility

10.6.1 Introduction

Normal operation of a base-station requires that careful attention be given to the impact of the high level transmitter signals on the very sensitive receivers which are co-located. This is largely covered by the overall equipment performance aspects which are covered by the Regulatory Authority such as the MPT series of specifications. The case under study here is however more vulnerable to direct and interactive effects due mainly to the close proximity of the 140 MHz receive bands to those of the transmitters in the 150 MHz bands.

It is necessary, therefore, to give detailed consideration to the extra requirements in this case, which manifest themselves as filters for receiver inputs and transmitter outputs. These filters will be additional to those likely to be already inherently contained in the associated active inputs. A diagrammatic representation of the base-station arrangement appears as in figure 10.6.1. The multichannel transmitter is to be considered as an alternative to the two 150 MHz transmitters shown in the diagram. The lettered boxes represent filters the specification for which will be derived in the next sub-section.

This work was necessary in order to produce (at least outline) specifications for all equipment early in the WARC programme. But the magnitude of the passive intermodulation effect was not known, nor was the detailed frequency plan. In fact it was not even known whether, if the passive effect were found to be severe, it would be possible to produce a frequency plan which would ensure that intermodulation products upto order n would be clear of receiver frequencies on that site (except for n=3 of course). Therefore two stages of derivation were undertaken, the first assuming 5th order products would be taken into account by the frequency plan (and therefore only 7th order needed consideration here), and the second on the assumption that 5th order needed consideration here.

10.6.2 Assumptions and Parameters

To determine the need for filters and associated devices which will be connected to the RF side of fixed transmitters and receivers, and to derive their electrical performance characteristics, three types of hill top configuration were considered:-

i. one transmitter per channel, one aerial per transmitter

ii. one transmitter per channel, one common transmit aerial

iii. multichannel transmitter, common aerial.

For each configuration two cases were taken into account

a. frequency assignments such that transmitter 5th order intermodulation products need not be considered

b. transmitter 5th order intermodulation products were considered.

Also

i. The frequency bands considered are:-

Main Transmit 70.5 - 71.5 (Fire), 152 - 156 (Police)

Main Receive 80 - 84 (Fire), 143 - 148 (Police)

Whilst there will be only partial occupancy of some of these bands, it was thought not worth sub dividing for the purposes of this exercise.

ii. Transmit radiated powers are equivalent to 50W per channel from an isolated dipole. Minimum receiver levels are equivalent to 2 μV emf (-137 dBW) into a 50Ω receiver, and that this level should be protected by 30 dB from each source of degradation.

iii. Losses from transmitter to aerial may total 3 dB, therefore single channel transmitter power will be 100W; losses from aerial to receiver shall be minimised (2 dB?).

iv. The aerial couplings/isolation are

a. between 150 MHz transmitters 20 dB

b. between 150 and 140 MHz aerials 30 dB

c. between 150 and 70 MHz aerials 20 dB

d. between 70 and 80 MHz aerials 30 dB

v. The intermodulation products will be contained by arrangements of filters associated with the transmitters and receivers as shown in figure 10.6.1.

vi. The total interference level is to be 20 dB below the minimum receiver level ie - 157 dBW, and any one contributor shall be 30 dB below minimum receive level ie - 167 dBW. This allows for a number of contributory interfering mechanisms to add their contributions (with the assumption of power addition) without exceeding the overall target.

10.6.3 Calculations - 140 MHz Receiver, Blocking

Normal receivers have their blocking level defined by MPT (Regulatory Authority) specification as the point where an unwanted signal causes what would otherwise be a 12 dB SINAD output, to fall to 6 dB SINAD, the input level for this is specified as 90 dB relative to 1 μV (90 dB μV).

The choice of the SINAD value is considered unfortunate for the users considered here, since at 12 dB SINAD wanted signals are unintelligible, at least to untrained users. Therefore the starting figure for the determination of the specification of the 'A' filter of figure 10.6.1 is that blocking will be deemed to occur at 10 dB better values for this. Additionally a 6 dB allowance is made to cater for more than one transmitter causing blocking. Thus the allowable level of any one interfering source is taken as 74 dB μV emf or - 69 dBW.

Now the level of any one transmission coupled into the receiving aerial is +17 - 30 = -13 dBW. Therefore filter A requires an attenuation of 56 dB at the transmitter frequencies. Its in band loss of 1 dB has been allowed for in the other calculations.

As discussed in section 10.5 the intermodulation performance of the receiver should be specified when operating with the filter attached. That is, if signals at 152 MHz and 154 MHz, each having an amplitude of

-13 dBW, are connected to the input of the filter then the effect on the receiver is to be no greater than a carrier of amplitude -167 dB at a frequency of -148 MHz. Alternatively, if the receiver is tested without the filter connected as above then the two carriers should be at a level of -69 dBW.

10.6.4 Calculations -150 MHz Transmitter (7th Order)

For this calculation it is assumed that frequency assignment will take account of all 5th order products leaving only 7th order ones to be dealt with here.

The filter which will result is designated type B.

The level of a product which can be tolerated at a receiver is -167 dBW which, with the 30 dB allowance for transmit to receive aerial coupling, means that the maximum level from a transmitting aerial is -137 dBW. Now sub-section 10.5.2 showed that 7th order products will have a dB slope of 3, ie the product level can be represented as 3T+K in dBW where T is the level in dBW of the signal coupled from another transmitter and K is a constant. From the measurements of section 10.5 and allowing for a 20 dB coupling loss between transmitter aerials, then the product level can be expressed as P7 = -3L -52 -R where P7 is the seventh order level in dBW, L is the selective loss or attenuation at the frequency of the unwanted transmitter and R is attenuation at the frequency of the product. It will be seen that every dB of attenuation which affects the level of the coupled unwanted transmitter (ie at the frequency of the second transmitter) is three times as effective as attenuation at the frequency of the product.

If selective filtering were used, ie selective in the sense of attenuating the unwanted signal from another transmitter but not attenuating the wanted transmission, then the difference in attenuation between these frequencies would have to be 19 dB. This assumes that the attenuation at the product frequency would be 30 dB as could be obtained for a single high Q cavity filter of the type described previously (the beer barrel). This 19 dB, is greater than could reliably be obtained from such a filter, so two would have to be used in cascade and the minimum separation between transmit frequencies would be 300 kHz. This is unacceptable, particularly when the problems of having to achieve the alignments on site are considered. An alternative could be an isolator, but again two would be needed in cascade to achieve the required attenuation (29 dB since R=0) and there would be inadequate performance in respect of transmitter noise.

The solution therefore is to use a single isolator of, say 20 dB isolation, followed by a bandpass filter having an attenuation of 25 dB at receive frequencies. But see the next sub section which will increase this to 28 dB. It will be recalled that this isolator is effectively called for in the transmitter specification.

10.6.5 Calculations - 150 MHz Transmitter (5th Order)

When considering the case where 5th order products must be dealt with, then the permitted level of such a product remains -137 dBW at the transmitting aerial, but the relationship between the selective attenuation, L in dB, and the product output at the aerial is $P5 = -2L -29 -R$. Therefore 2L+R needs to be 108 dB. Two solutions are possible; either retain the 20 dB isolator and make the rejection of the type C filter required here to be 68 dB at receive frequencies, or to leave the filter as an enhanced type B, with 28 dB of rejection, and insert an additional 20 dB of isolation. In either case the noise specification is met and the second option is very attractive since only one filter type need be specified and the second isolator can be fitted retrospectively if at a later stage it is found necessary to cater for 5th order products. [As a note of warning this is all in advance of the detailed consideration on passive intermods of the next chapter].

10.6.6 Calculations -70/80 MHz Transmitter and Receiver

Frequencies in the band 70.5 to 71.5 MHz from fire service transmitters will interact with those in the 152-156 MHz band to produce 2nd order products in the range 80.5 to 85.5 MHz which encompass the fire service receive bands. The conversion loss for this straightforward mixing process could be as bad as 6 dB. To bring the products to an acceptable level at the 80 MHz receivers then the 150 MHz filters must have the sum of their attenuations at 80 and 70 MHz to be at least 118 dB. The distribution of this sum for the filters, which have been designated as B and C, could be sensibly 55 dB at 80 MHz and 65 dB at

70 MHz. [The isolator(s) cannot really be relied on to perform this far from their designed (150 MHz) operating frequency] these figures therefore define the skirt responses of filters B and C.

An equivalent filter will be required in the 70 MHz transmitter output - filter D. Here, the sum of attenuations of 118 dB, applies to the 80 and 150 MHz bands and, although the 50 MHz band is closer to the pass-band than in the 150 MHz filter case, the narrower pass band will allow a moderate attenuation to be planned for at 80 MHz. Thus a sensible split would seem to be 40 dB at 80 MHz and 90 at 150 MHz. The determining factor here being the difficulty of establishing actual filter performance when more than 90 dB is specified.

When the 80 MHz receiver filter is considered, the stop band attenuation requirement is the same as that for the 140 MHz receiver filter and this then determines the specification for filter E.

10.6.7 Calculations - 150 MHz Transmitters and Combiners

If the combiner input port isolation is taken as 20 dB, and it is assumed that the aerial gains in this case are adequate, and with the same couplings, then the calculations for the non combiner case are repeated. Thus the specifications of filters B and C remain the same.

Another note of warning here is that aerial mismatch will effectively degrade the input port isolation of the combiner. The maximum expected mismatch of the aerials has been allowed for and the combiner specification requires the 20 dB to be obtained with all degrees of mismatch upto this level.

10.6.8 Calculations - 150 MHz Mutlichannel Transmitter

Since significantly more channels will be radiated from the transmitting aerials in this case compared to the single transmitter per aerial situation, then more attention must be paid to receiver blocking aspects. Rather than call for a different receiver filter, the best solution appears to be to call for better transmit to receive aerial isolation in this case. A figure of 40 dB would suffice and is in fact achievable with the skeleton slot aerials described in sub-section 10.2.4.

A filter will be required in the transmitter output to bring the 7th order products to an acceptable level for the receiver. Such a filter would reduce the specified level of 70 dB below carrier (-57 dBW) to -167 dBW at the receiver input. Thus a requirement of 70 dB attenation at receiver frequencies is called for:- this is the requirement for filter F.

Similar consideration apply when 5th order products are assessed, and the corresponding filter attenuation at receiver frequencies is 75 dB:- this is then the requirement for filter type G. It is a moot point as to whether it is sensible to continue with two options which are so close in specification. In addition the filter can be effectively incorporated in the multichannel transmitter by including the required performance in the specification (it is in any case a special purchase item so that the specification can be tailored to suit).

10.6.9 Summary of Filter Specifications

The filter performance requirements are summarised in table 10.6.1 below.

Table 10.6.1   Filter Requirements

Filter type	Notes	Use	Minimum Attenuation dB In Frequency Bands - MHz				
			70.71.5	80-84	143-148	152-156	162.5
A		Police Receivers	60	-	1 (Max)	56	
B	1	Police Transmitters excluding 5th order considerations	90	40	28	2 (Max)	
C	1, 2	Police Transmitters including 5th order considerations	90	40	68	2 (Max)	
D	1	Fire Transmitter	1 (Max)	40	60	90	
E		Fire Receiver	56	1 (Max)	-	60	
F	3	Multi channel Transmitter excluding 5th order considerations	90	40	70	1 (Max)	
G	3	Mutlichannel transmitter including 5th order considerations	90	40	75	1 (Max)	

NOTES:

1. These filters assume the presence of a 20 dB isolator in the output of the relevant transmitter. It is believe that this will have to be incorporated in the transmitter to meet the specification requirements in terms of intermodulation performance.

2. This filter could be replaced by one of type B and an (additional?) isolator of 20 dB isolation.

3. A filter which will meet the spurious/intermodulation performance to this standard may be called for as an option in the multichannel transmitter specification.

The table (10.6.1) provides the essential data for a filter peformance mask. Such a mask is shown in figure 10.6.2 for the filter type B by way of an example.

10.7 Conclusions

This chapter has examined aspects of base-station design following largely the elements shown on the interaction diagram. In effect many passes around this design loop have been made and a number of option proposed.

Whilst some comments have beenmade on the influence of other design aspects these have been kept to a minimum to try to aid clarity. It is after all intended that more traverses of the interaction diagram be made before reconsidering the base-station design. A particular element which forms part of the base-station, but has received little comment here, is that of the fixed links. This forms part of the fixed link design loop also, but little change is foreseen here unless forced. The main need will be to include suitable transmit and receive filters and the specifications for these could be derived in a manner similar to that of the last section.

11. Intermodulation - Site Non Linearities

11.1 Introduction

11.1.1 General

Descriptions have been given in chapters 9 and 10 of two of the three places in the mobile communications system where intermodulation products arise. These were the mobile receiver, and active elements of the base station (transmitters and receivers). For the second of these it was shown that the impact on system performance could be contained by suitable design, mainly the aerial dispositions, filtering and isolation. For the first situation the phenomenon will only be manifest in a region around a base station and is usually accommodated in the overall design by suitable frequency assignment strategies (see chapter 12).

The third location for the generation of the intermodulation product (intermod) is that of the passive base station elements such as feeders, filters, aerials, towers - and the site in general. Since the non linearity involved is thought to be due to imperfect contact between jointed metals, which has been exacerbated by corrosion, the term 'rusty bolt effect' is used as a descriptive cover. It must be emphasised however that is is not necessary to have a bolted structure or for there to be actual rust present for the effect to be present. In fact the junction between any two metals is suspect and attention has been drawn to surface and bulk non-linearities (such as magnetic hysteresis) and dielectric non-linearities as causes.

The degree of non linearity is not large and therefore the products are at low level. Therefore the influence of such intermods at any significant distance from the site is usually very small. But they obviously will affect any sensitive equipment on the site; and the most sensitive are the receivers.

Thus the passive intermods differ from the other two forms in that their expected levels are very low, but they also differ in the respect that once generated on a receiver frequency no filtering can be used to prevent the intermod entering the receiver. All that can be done is to try to isolate the receiver from such products (this chapter will show that this cannot be done to the degree desired), and/or to produce a frequency plan and assignment strategy which steers receiver assignments clear of the intermod, and chapters 12 and 13 will show the difficulty of doing this.

11.1.2 Initial Factors

Appendix J forms a brief tutorial on intermodulation theory. It is derived from a contribution by the author (Gardiner and Fudge 1984) to a vacation school on mobile radio, and covers the manner in which intermod can be considered to be produced. It also defines some of the terminology in general use, such as intermod order number - n.

Chapter 8 considered the influence of propagation factors on the choice of the frequency bands for use in the post WARC situation, it also touched on the note of caution and uncertainty which was sounded to those negotiating the new frequency bands for the aspect of site generated intermods. The evidence at that stage was rather flimsey. As the result of some exploratory work two measurements had been made of 5th order site products. These are shown in figure 11.1.1. The two points shown on that graph were measured at different times at different locations and with different powers. The only jutification for putting them on the same graph is that they were the only measurements available at that time for the Directorates cases, and a line of the expected slope connected the two. They were made under conditions of great difficulty since the apparatus seemed to be not completely stable - we now know that this is in fact due to to the nature of the phenomenon itself. The indications were however that 5th order products if they were to fall on receiver frequencies would be damaging and at that time it was not known whether a frequency plan could be produced which would avoid such products.

The previous chapter on base station design, derived the target figures for any one contributory cause to the receiver's overall interference as being 30dB below the minimum usable input level of 2μVemf (-137dBW). To measure intermod products down to the desired level of - 167dBW would require a measuring apparatus set capable of working to this level, in itself a difficult task, whilst nearby carriers of amplitude

50W (+17dBW) were present. Thus an overall range for the test set of +17 to -167dBW, or 184dB was called for. This enormous range is also an indication of the tolerance (or perhaps the intolerance) of the communication system to any non linearity.

11.2 Apparatus for Measurements

11.2.1 Power Sources

The measuring apparatus comprised of three main sections:- an intermodulation free source of carriers whose amplitudes will be varied, a means of coupling them to the system under test (whether filters or towers), and an extremely sensitive (and intermod free) receiver. An overall block diagram is shown in figure 11.2.1.

To obtain the power required, over 100W, two linear power amplifiers were used. They were driven by signal generators of high purity and low noise, several types of high quality generator were used, but for maximum versitility the same devices as used for the modulation bench tests were employed. These had the very high frequency stability required but with easy frequency changing, accurate power control and variability, and above all the ability to be programmable for intended future automation.

To prevent the generation of intermodulation products in the output of the power amplifiers a cascade of cavity filters preceded by an isolator (circulator plus load) described in section 10.3 were used. The situation is akin to that of the design of the base station couplers, and transmitter outputs, as discussed in chapter 10 but there is not the absolute need here for the two transmitters to be very close in frequency so a somewhat different solution can be adopted. Power sensors were included in the final output in order to monitor the power but care had to be exercised here.

The type of power sensor generally used was that known as a 'Bird Thru Line'. This consists of a short section of coaxial line which is inserted in series with the feeder to be measured. In the inserted device is a very small coupling loop which senses the power travelling in one direction and displays it on a meter after rectification by a diode detector. Although the coupling is very low, the diode was found to be a cause of intermods - not unexpectedly. It was necessary therefore to remove the coupling loop from the power monitor, which was a design feature of the power meter, before making intermod measurements. Alternatively the output of a similar coupling loop without detection could be taken to a thermal power meter. This provided a more accurate measurement but again it was thought unwise to leave it in place during measurements.

11.2.2 Connection Devices

If the device under test in a two port device such as a filter, feeder or an aerial then the two transmitter sources are coupled in the manner shown in figure 11.2.2. At the junction each input branch appears as an open circuit to the frequency of the other source. This is achieved by the use of critical line length approximately half a wavelength long. Figure 11.2.2 shows the arrangement for the receiving side when testing a two port device. The length L is critical to make the receiver branch appear as an open circuit off its tuned frequency. To calibrate the test set the two port device under test is removed.

If two separate transmit aerials are to be used to explore the tower then the coupling between the aerials is simulated by the arrangement shown in figure 11.2.3. Here C is a length of good quality (non braided) coaxial cable of sufficient length (equivalent to 20 dB loss) to prevent any intermodulation products of the resistive load L from being noticeable. It must be emphasised here that all components should be viewed with a jaundiced eye as far as their intermodulation performance is concerned, resistive loads driven at full power are not acceptable in this respect. The C_A's and C_B represent other cable loads chosen to simulate the transmit and receive aerial couplings, see appendix J. The length L is set so that the receiver branch appears as an open circuit at the frequencies of the transmitters. Ideally C should be preceded by a filter rejecting the frequency of the intermods so that the receiver junction forms a matched diplexer, but the inaccuracy in the previous method is not significant compared to the overall accuracy and the difficulty of achieving a good enough intermod performance from that filter.

11.2.3 Receiver

The function of the receiver is to measure the level of external intermodulation products in the presence of high level signals which will cause blocking and generate intermods in the receiver's front end.

The blocking can be overcome by suitable pre-filtering and this will also reduce the level of receiver generated intermods, even so the requirements are severe. The requirement is for high dynamic range coupled with high sensitivity. The spectrum analyser field was assessed but could not meet either of the requirements. Communications receivers showed promise on one or other of the counts. In the end a very high quality receiver was selected which had over 90dB of dynamic range and with an input stage of very high linearity. This, whilst sensitive, could not resolve the lowest level signals due to its narrowest bandwidth (15kHz) passing too much noise. Since fixed frequency signals were being measured then in principle inserting a narrow filter (one hundredth of the bandwidth would yield the required 20dB extension) would suffice. But firstly there were difficulties in constructing such a narrow filter at the receiver's intermediate frequency, and secondly the receiver's frequency stability was not sufficient to use it.

Instead the IF output was connected to a high performance spectrum analyser which both covered the IF frequency range and had a 50Hz filter. Using the spectrum analyser in a sweeping mode enabled the system to cope with frequency uncertanties up to the sweep bandwidth.

All units were inherently digitally controlled and hence programmable, which was viewed as a likely future requirement for automation.

11.2.4 Summary

The equipment used was:-

Signal Generators	-	Marconi Instruments type 2020
Power Amplifiers	-	ENI type 3200L
Isolators	-	Phelps Dodge Circulators with resistive load
Power Meters	-	Bird Thruline
Cavities	-	Aerial Facilities
Receiver	-	Rhode and Schwartz type ESU2
Spectrum Analyser	-	Marconi Instruments type 2370

11.3 Measurements

11.3.1 Basic Measurements I

The initial aim was to simulate conditions which were to be expected in the post WARC system, and measure the level of intermodulation products. From this it should be possible to guide the frequency planning exercise as to whether 5th or even 7th order products were likely to be troublesome. Therefore the first tests were performed on a tower which was convenient for the laboratory. It was in fact located in the inner courtyard of the square 3 storey laboratory buildings, but its height projected well above the roof level.

Three dipole aerials were deployed with the two transmit ones spaced to give 20dB coupling between them, and the receiver one placed symmetrically below these with 30dB coupling to either. Thus the aerial equivalent of the appendix J system was achieved.

By a suitable choice of transmitter and receiver frequencies (and the retuning of the associated filters) it was possible to measure the levels shown below. The intermodulation product level is referred to the level of the minimum wanted signal - 137 dBW. Figure 11.3.1 shows a compendium of results obtained over a period of time and at different locations, they are presented in this form to provide a saccinct summary in one location, but they will be described in sections and the section under examination here is designated A in the table.

It will be seen that the Harrow measurement covered two weather conditions; dry and wet. These very first results indicated that the intermod levels were worse in wet conditions than in dry - subsequent measurement here and elsewhere reversed this finding. However two main points were immediately apparent from these results. Firstly the levels were higher than expected and might be troublesome up to 11th order. Secondly the measured levels were very variable with excursions of 30dB or more, so that the figures given represent some kind of average for the best and worst five minute periods over a total of half an hour for each. These results were rather disturbing, therefore, in showing that the passive intermod phenomenon not only was non-negligible but in fact was more extensive than feared.

Since the Harrow location was in an industrial area where there could well be other influences, it was decided to repeat the measurement at one of the Directorate's maintenance depots at Cheveley where there was a suitable tower which could be cleared of all other users for the period of experiments. The results are shown as section B. The first line of these results show quasi peak levels since these are thought to be more meaningful. They will be seen to be even worse than the Harrow results here even 15th order products would be troublesome! Thoughts were expressed that the non-linearity could be either in the transmitting aerials themselves or very close to it. Therefore moving the receiving aerial further away should be beneficial. The effect of this is shown in the second line of B where the receiving aerial was mounted very low down on the tower such that its coupling to the transmit aerials was 50dB. This made no significant difference! The weather conditions here were dry and the variability in level was attributed to tower movement cause by the wind/differential solar heating. The intermodulation levels could be affected by hitting the base of the tower with a bulk of timber or by pulling on a rope attached to the top of the tower.

One of the other disturbing features here was that the Cheevely tower was only six months old (and of the current standard of galvanised angle iron construction) whereas the Harrow tower was some 15 years old (galvanised tubular construction).

11.3.2 Basic Measurement II

Further measurements were conducted at Cheveley with other aerial postions for both transmitters and the receiver. The results of these are not shown in the figure, but they confirmed that the intermodulation level was little influenced by any of the aerial coupling figures involved. This, then, indicated that the source of the intermod was not in the aerial itself. It was as if the tower was aglow with intermods from top to bottom and a guided wave phenomenon, supported by the tower, was postulated.

At this stage it was possible to take advantage of the aerial evaluations described in section 10.2, and an array of skeleton slot aerials were deployed on the four faces of the tower. These were of all-welded construction and it was thought that the associated reflector elements would shield the tower from the high levels of transmitter signals on the outgoing side, and on the incoming side shield the receiver from the tower generated products.

This series of measurements, still at Cheveley, is shown as section C. The isolation between any two skeleton slot aerials was approximately 40dB and the first tests were conducted with the transmitters each coupled to its slot aerial and the receiver connected to a dipole in various locations on the tower (Tx - Rx isolations 42 - 60dB). The levels of intermod products are certainly lower but not very much so.

The receiver was then connected to one of the slot aerials and the intermod levels again improved. Most surprisingly the results further improved when the two transmitters were coupled to the same one slot aerial. All these results require caution in their interpretation, since the large variability was present, and subsequent measurements were to show that there could be diurnal variations.

Some comments can be ventured however; if the intermods were ocurring in one of the transmitter aerials then the change of some 20dB in their isolation would be expected to result in a very significant change in the intermod level. Appendix J indicates that, for say the 9th order case, the variation should be at least 80dB assuming a polynomial representation. This is certainly not the case. If the non-linearity, were to be in the receiving aerial then changing the coupling of this to the transmitters should result in even larger changes in intermodulation products (20dB change in isolation should give 180dB change intermod for 9th Order) thus it would seem that the aerials can be exonerated.

The same sort of reasoning applies to a tower joint if it is considered to be close to one of the aerials first discussed.

Overall this series of measurements showed that even in the seemingly ideal case where skeleton slot aerials were to be used for both transmit and receive (two separate arrays) then intermodulation products upto at least 11th order should be taken into account when producing a frequency plan. This carries serious implications since such aerial arrays will only be used where there are a large number of channels and the spread of the intermods will be extensive, as will be discussed in section 11.4 on implications.

11.3.3 Basic Measurements III

Thus far only two sites had been measured and, although they were of different ages and construction, there was no indication that they were representative of those to be found in use. They were after all both normally non operational towers and had electrical activity in the vicinity (either the laboratory or a depot). It was therefore decided to take the opportunity of measuring an operational site. Normally this would be very difficult since all the operational transmitters and receivers would be coupled to their aerials on the tower and these active devices could form the sources of intermods unless they were heavily filtered.

A site at Ousden was undergoing a tower replacement programme and access to the new system was available before the aerials were to be changed from the old towers to the new. Furthermore it was a two tower site. But a further note of some caution must be sounded in interpreting the results since the old towers were only some tens of feet away from the new.

The results are shown in figure 11.3.1 under section D, the first line relates to the normal dipole aerial disposition on one of the new towers. It will be seen that the levels are a little lower than the equivalent situation at Cheveley but the significance of the difference is uncertain in view of the variability. When the receiving aerial was transferred to the other tower the intermod levels fell markedly but not as much as might have been expected - certainly for the higher orders. The most dramatic improvement was obtained by placing one transmitter and the receiver on one tower and the other transmitter on the other tower. The product levels were then at about the limit of sensitivity for the equipment for orders 5 and 7, and below this for higher orders.

This indicates the desirability of retaining the two-tower operation of sites wherever possibe, but splitting the transmissions between the towers, and similarly the receivers, instead of the current practice of assigning one tower to a transmitting function and one to receiving.

The results from this third site were very much in line with the other two. Throughout this series of measurements the reorganisation of the

Directorate was under way and there were many staff movements. It was difficult to impress on others involved in the WARC planning that the implications of these results were very serious, particulary for the frequency planning aspects. It was therefore decreed that further work on this topic would be undertaken by placing contracts with Universities and Industry. They would both investigate the details of the phenomenon, and the extent to which it occurred nationwide, and hopefully to find a cure. Due to these factors, therefore, only sporadic measurements were carried out in-house and this will be discussed in the next sub sections. But it was impossible to bring the work to a sensible conclusion - virtually all available effort went into the management of the contacts which, at the time of writing, have yet to be started due to contractual and financial difficulties.

11.3.4 Variation with Aerial Position

It seemed obvious at first that even the passive intermod effect would decrease as the distances between, and hence the isolations of, the aerials increased. Whilst this appeared to be true from the results, the degree of variation was far less than expected. The theory was advanced that there might be standing waves (of either the fundamentals, or the intermods) on the tower or that the intermod levels could be greatest around concentrations of joints such as occur at set intervals in many towers. The results shown so far have only included some crude adjustment of receiving aerial position.

It was therefore decided to explore the intermod level variation with position, in height terms, of the receiving aerial. Ideally this should have been moved up the outside of the tower with constant spacing from it. This however was impractical since it would involve an aerial rigger having to continuously reposition the aerial. Instead a receiving probe was constructed in the form of a wire aerial in the shape of an H, the upper wire taking the form of a rectangular U and the lower one being an inverted version of that; the length of the wire was half a wavelength and the feeder was connected to the centre points of the horizontal arms. The wire aerial was mounted on a rectangular wooden board and this could be hauled up the centre of the tower by nylon ropes and was constrained by running on nylon guide ropes.

Thus a sensing aerial (of unknown, but constant, polar response, and moderate match to the feeder) could be easily and quickly postioned at any desired height: the feeder was light-weight flexible (braided) type.

Again time and resources permitted only a single session of measurements of this situation at Harrow and the results are shown in figures 11.3.2 and 11.3.3. These are for 3rd order products using two different tansmitting aerial arrangements.

There is some general increase in level around the tranmitting aerials as might be expected due to the high field strength, but there are other locations of high strength which show no real periodicity nor correspond to any features of the tower. In figure 11.3.2 there is a general fall in level from the 50ft height, which corresponds roughly to the height of the surrounding laboratory building, but this is not repeated on the other runs. The receiving aerial could be considered to be shielded from the normal intermods since it was located in the centre of the tower whereas the transmitting ones are on the outside. But this is unlikely to be the case with such a open structure and was not the case when other dipoles were located on various faces of the tower.

The spread of reading at each height is shown and it is noticable that there is little, if any, spread for heights below the surrounding building. The measurements were performed on dry cold days with little wind, but even this wind would move the receiving aerial noticably sideways on its rope suspension. Therefore the variability shown is more likely to be due to this movement than any wind induced tower movements.

11.3.5 Variation with Transmitter Power

The behaviour of the passive intermodulation products thus far were not consistent with any simple model. They were however significantly higher than could be tolerated, and if aerial isolation, either between transmitters or transmit to receive, could not effect an amelioration then it was possible that transmitter power reduction could produce the desired result. Appendix J shows the traditional expectation for order n of an n:1 decrease in intermod level when both input and output are measured in dB. For 11th order therefore a 3dB reduction in tranmitter power should cause a 33dB reduction in the intermod level - quite adequate.

Some tests were therefore carried out to determine the variation with power level.

Figures 11.3.4, 11.3.5, 11.3.6 and 11.3.7 are derived from measurement by R Keeble (1981). All were measured at Harrow and the results of figure 11.3.4 are for skeleton slot aerials for transmit and receive. It will be seen that good straight line relationships were achieved but the slopes were less than expected - increasingly so as the product order increased. The conditions were dry.

Figure 11.3.5 shows the results for two different receiving aerials. The transmitting aerials were again slots at 140ft level and the upper curve was for the receive aerial being a slot also at 140ft. The lower curve is for a folded dipole at 65ft. Again dry weather. The slopes in this case (3rd order) were nearer that expected; but they were in close agreement with each other not only on slope, but also on absolute levels, which is perhaps surprising for so different a form of sensors and more particularly for the different heights.

The curves of figure 11.2.6 are those for a different aerial arrangement. This used the folded dipoles described earlier in this chapter as simulating an expected operational condition. They thus represent the tight coupling conditions. The upper curve was taken in dry conditions and the lower in wet. Here there is a marked difference in slope with the higher levels of intermods being apparently quenched by the rain. There is an indication of reversal of trends at very low levels however.

Figure 11.3.7 shows the effect of various locations of the receiving folded dipole aerial for the same transmit aerials as just described. The gross changes in height, or which leg of the tower they are attached to, makes little difference, the slopes are the same and all significantly less than 3.

The curves of figure 11.3.8 were obtained for measurements of a new (9 months old) tower, purpose-built at Sandridge. They are for a 7th order product taken in dry but very humid and cold conditions, with probable tower icing, of zero wind; and very stable results were obtained. A degree of hystersis is shown, but this may well be another manifestation of the general variability. The discontinuity at high power levels could be explained by equipment deficiencies.

The change from a straight line occurred at the point where the level setting of the signal source of figure 11.2.1 changed from one of just over 100dB (102.5dB) to one of just under (99.5dB) with the apparent explanation that 100dB was switched out on the attenuator and 90+ dB put in. Tracking of these two, to the high degree required for these tests, would be difficult. But several factors tell against this explanation. The first is that the attenuation is presented as a decimal coded switch which remotely controls the setting of a modulo 2 attenuator cascade. Therefore a change from above to under 100 does not necessarily mean a large change in the attenuator settings. Secondly the term attenuation has been used, but the presentation is the inverse of this; it is level (dB wrt 1μV) and so the actual attenuation used depends on the internal highest power level. Thirdly the power levels were subsequently checked at the output of the filter chain and found to be sustantially correct throughout the range.

The high level discontinuity therefore remains to be explained. The slope of the graph is however very constant at 3.3, which is very significantly below the value of 7 expected.

Overall it can be said that the intermod levels do fall with falling transmitter powers but the rate of fall is nowhere near as great as might be expected from simple polynomial theory predictions. Therefore to achieve, say a 30dB reduction in 9th or 11th order product levels, the decrease in tranmitter power would be unacceptable on range considerations.

11.3.6 Variations with Time

Several comments have been made regarding the apparent variability of the intermod levels with time. Some of the results shown have indicated this as a spread of values. Work was carried out under contract to the Directorate by the University of Southampton on the prediction of intermod frequencies and magnitudes. The measuring set already described was adapted by them to drive a chart-pen recorder. This was used to record the levels of intermods against time.

Some of their results (Debney and Stewart 1983) are reproduced here as figures 11.3.a, 11.3.10 and 11.3.11.

The first of these shows the variations of a fith order product over an hour around midday on a dry and sunny day, using dipoles. The constant horizontal line is a measurement of the amplifier output power, and with an overall chart sensitivity for this of 5dB it will be seen that the exciting signal was constant. Turning to the intermod chart it will be seen that the initial level was moderately low (-125dBW) and steady until affected by external interference. After a slow climb to some -110dBW there were a few rapid flunctuations and then 25 minutes of very little variation. The second half of the plot shows the more normal rapid variations with amplitudes of 20dB.

The second figure (11.3.10) was taken later that same day. The overall variations are greater with excursions reaching -100dBW, which in 37dB above the minimum expected wanted signal and 67dB above the design allowance. This plot is typical of normal measurements. For figure 11.3.11 attempts are made to determine the correlation between the levels for two different orders. The trace shows therefore both 3rd and 5th orders, the scale is correct for the latter and should be read as 20dB higher for the former. The most noticeable feature is the extreme constancy of the levels. This could be attributed more to the calm wet condition than the change to skeleton slot aerials, although this is of course conjecture. Two day's traces are shown, the second day exhibits a change in the relative level of the 3rd and 5th but since there were virtually no observed variations (less than 2dB) the connection between the orders was unproven.

Overall it will be seen that conditions such as those shown in figure 11.3.10 are not suitable for assessing intermod level variations when changing other parameters, such as tranmitter power etc. It is necessary to pick at least a calm day and possibly a wet calm day to get the abnormally quiet condition shown in figure 11.3.11. But the wetness factor may in fact condition the results obtained for these other parameter variations.

11.4 Discussion of Results

11.4.1 Initial

The major feature of the results of the measurements was their unexpectedly high values. When presented with unexpected results the experiemnter usually searches for faults in the apparatus or technique and this was done in this case. One telling experiment consisted of connecting both transmitter and the receiver to the welded skeleton slot aerial. This aerial was just laid on open ground. If it had its beam directed into the ground or at the sky then the intermod levels were very low. If however a piece of jointed metal was suspended in front of it, or the beam was directed at a metallic structure (tower or vehicle) then the intermod levels were very much higher. Thus all investigatory experiments to prove the apparatus and technique pointed to the results obtained being real ones.

Immediately the levels were found to be so high, then this was notified to others concerned with the WARC programme because of the extensive implications - these are discussed in the next section.

11.4.2 Intermediate

To-date fairly extensivemeasurements have been made at three sites and less extensive ones at other locations. Despite the range of tower types and ages the results are very similar. They exhibit intermod levels well above the minimum desired for up to at least 11th order products. The levels vary with time both in the short and long time scale. For the short ones it is thought that tower movements caused by wind or differential heating and cooling are likely causes. In the case of the longer term there is the dominant feature of weather again - less differential heating, and cold (still) winter days generally yield lower levels.

The weather also seems to be a dominant feature in its own right in terms of rain. If towers are saturated then levels are low.

The succeptibility to transitter power variation is not as great as would be predicted from simple theory. This might be another manifestation of the same feature which shows much less variation than expected in intermode level as aerial positions are changed. It is as if the tower was bathed in intermods.

There are no final results, those given here must be regarded as part of a continuing story, but enough has been done for this to be taken into account in system design.

11.4.3 Proposed Future Work

The implications of the high level of intermods fall largely in the area of frequency planning/assignment. The next section will show that they are severe and although chapters 12 and 13 will propose a strategy for tolerating the situation it does so only when all the emissions which could affect users at a site are under one control. This is frequently not the case. It was thought highly desirable therefore to have a programme of work which would hopefully provide a cure.

The programme would divide into three categories, developing an understanding of the phenomenon, assessing the extent of the occurrance, and discussing cures. All such activites must be viewed against the background of lack of manpower resources due to other committments.

To further the understanding of the nature of the effet there are plans to automate the measuring apparatus so that longer runs can be performed and machine analysis applied to the results. This was proposed by the University of Southampton as part of their contractual work, but will in fact only come to fruition outside that contract. It is proposed that the fundamental properties of intermodulation producation by metallic surfaces and imperfect joints be studied by The City University under a similar three-way contract of University-SERC-Directorate as that for Southampton University. There is also a proposed internal programme of constructing a sensing device to explore the tower structure to locate the actual sources.

Contractual negotiations are continuing with an industrial company to characterise a number of operational base station sites. This will provide an independant assessment of the levels and extent of the intermods. They also propose assessing cures which they have devised for other intermodulation production situations based on the observation of the quenching action of rain or water.

Other curative assessments are proposed by the Directorate which involve having a number of portions of towers (say 20ft top section) erected on suitable bases in the ground. Each section will be well separated from its nearest neighbour so that its intermod performance could be measured using the single external aerial technique described in sub section 11.4.1. The various sections would enable different materials and construction to be tried (including all-fibre glass) over a period of time and sinultaneously various cures assessed for immediate and long term effects.

11.4.4 Possible Explanations

The results presented seem to defy simple explanations of their causes, but, at the risk of being proven wrong by measurements yet to come, some comments are offered.

A single (point) source of the intermod seems to be ruled out by the small range of variation of level with position of either transmitting aerials or receiving ones. Traditionally it has been assumed that the metallic joints, being imperfect, are the cause. If so, and there is no reason here to doubt that, then there is no reason why any one joint should be better or worse than any other. Therefore all joints are potential sources.

This still does not however account for the small variation of level with locations. A source close to the transmitting aerials might be expected to produce strong intermods due to being located in a high signal strength area. However these intermodulation products should have fallen to a low level for an aerial mounted at some distance on the tower, due purely to that distance. If the sources were near the remote receiving aerial then the coupling to them would be stronger but the levels should be much lower in the first instance due to the exciting signals being lower. The phenomenon has the appearance of there being a guided wave (either the exciting signals or the intermods) supported by the tower structure itself. This guided wave then propagates up and down the tower with low loss, and accounts for the description of a tower aglow with intermods. It would also account for the effect noticed at Ousden where intermod reduction was only significant when one of the transmitters was removed to another tower. If only the receiver were removed then this could sense the 'glowing tower'.

A very brief experimnent was conducted to test this hypothesis by placing on the legs of the tower between the transmitting and receiving aerials, devices intended to attenuate any surface or travelling waves on the structure. The results were inconclusive - not all the legs were treated and they could additionally be by-passed by the cross braces. Another brief experiment did tend to confirm that the joints were the sources. This consisted of radiating two signals differing by 1MHz at 150MHz and searching the structure for the 1MHz difference frequency using a domstic portable radio tuned to the difference frequency. Strongest signal waves were reported around the structure joints.

If all the joints are equally likely to be excited, whether by the wave phenomenon or some other, then the net effect will be very similar to that of multipath described in the chapters on propagation. Each source may be efffectively on or off depending on its mechanical loading at any instant (which will change with wind or temperature movement), or even at various intermediate states. Thus any receiving device will expect to see a large number of signals having random amplitudes and, due to the different distances, random phases. This then would be expected to produce a field with level distributed in the Rayleigh manner of appendix D. Any proving of this would require that a statistical analysis of levels be conducted for various locations. The results of the trolley measurement of section 11.3.4 are insufficient for this.

It is traditional to view the tower as consisting of a number of dipoles (not necesarily resonant), with the joint at their centres. This is shown very diagrammatically in figure 11.4.1, where the D symbols are intended to represent the non linearity of the joints (for example a diode could be included in each joint). One outcome such a representation is to show the tower as a number of parasitic aerial elements, and this lends some support to the view that it will support the travelling wave and also account for the conduction of signals across the tower. A number of such parasitic elements would be in the near-field of the exciting aerials and therefore be driven by them - handing on the wave in turn.

The other outcome is that the non linearity sensed at any one receiving aerial location would be the sum of the diode characteristics with various (but constant for each) weighting coefficients. Therefore as the tranmitter powers were varied the Rayleigh field pattern would vary in amplitude only, not position, and any one location would be representative. Insufficient consideration has been given to the evaluation of the effective overall non-linearity characteristic of such a sum of individual ones. Whether this would account for the non-polynomial behaviour is untested but seems unlikely.

An alternative explanation for the non-polynomial behaviour has been advanced (Bailey et al. 1980) and it is shown that the joints of such a tower structure could support quantum tunnelling transfer characteristics. Again this aspect has not been followed up.

But the "diode tower" structure seems to ignore the open ends of the dipoles. These will in fact be joined to the open end of other dipoles in a real structure, and this is shown, again very diagrammatically, in figure 11.4.2. Alternate points are here labled D and S. The D is again to be understood as a non-linear joint or diode, but the S can be taken to refer to sparking. If the structure were really as shown in figure 11.4.2 and the diodes were resonant half waves, then the D's could be taken as relatively low impedence joints and the S's as higher impedance ones. Then when the diodes were excited there would be a high potential across the higher impedance since the dipole ends will be of opposite potentials. It is feasible therefore that break-down could occur in the form of arcing or micro sparks. This again is a non-linear phenomenon which could perhaps account for the overall non-linearity as much as the diode pheonomenon could, and would additionally give rise to a non-polynomial characteristic. It could also provide a better explanation for the effect of water on the joints in that its presence would obviously modify the joint impedence markedly and could well suppress the sparking.

The explanations at this stage must however be considered speculative, but worthy perhaps of being taken as postulates for the further work necessary in this very interesting field.

11.5 Implications

11.5.1 General

The intermodulation products produced by the active elements of a base station were shown to be capable of being brought to an acceptably low level by the use of filters and isolators. The passive intermods are, however, in an entirely different category in that the only way of stopping them from causing interference to colocated receivers is to prevent their production by ensuring that there are no non-linearities, or to arrange for suitable frequency assignments which avoid intermodulation product frequencies falling on receiver frequencies.

This perhaps requires some explanation. Since the source of the intermod is assumed to lie electrically beyond the feeder/transmit aerial junction, then no frequency selective filter, or non reciprocal device, can be incorporated into the transmit side to attenuate an unwanted emission. Therefore the level of the generated intermod is dependant on the non-linearity itself and the coupling is largely independent of aerial positon. Once generated, then if an intermod falls on a receive frequency, no receiver filtering or non-reciprocal device can be incorporated to attenuate the intermod without similarily attenuating the wanted receive signal. Again, the only avenue open is to decouple the receiving aerial from the source of the intermod, and although the use of aerials with built in screens, such as the skeleton slot, gives some advantage it is not enough.

The measured levels indicate that products of orders up to and including the 11th are above, and frequently well above, the maximum desired level of - 167dBW, where they could contribute significantly to the background noise level when receiving a weak wanted signal. Frequently the levels are above the receiver mute setting which means that the listener, the controller, will experience interference when there is no incoming message.

The true implication of the unexpectedly high level of high order intermods will form a major topic in the chapters on frequency planning and assignment (12 and 13) where the interaction of other factors will be brought in. It is also to some extent covered in appendix J, but this chapter would be incomplete without a description of the major impact which was realised as soon as the first results of the Harrow measurements were available.

11.5.2 Broad Details

In advance of knowing detailed frequency assignments in the transmit bands to be used, as described in section 8.5, it was not unreasonable to assume that at any base station radiation could occur anywhere in the transmit bands. Since both mobile and fixed channels were to be accommodated in these bands then there would be many emissions at any site and a large number for the high usage sites. It was reasonable to assume that they could be spread anywhere within the assigned frequency bands and therefore a global, or broad-brush, approach suggested that the transmit bands be considered as a continuum of tranmissions, this was in conformity with the view that emissions would be spread widely in frequency, rather that clumped together, due to the desire to avoid 3rd order products being generated in the mobile.

This then leads to the production of a very simple picture of the extent of intermodes of various orders and is shown as figure 11.5.1.

As was realised at the time of discussion on the choice of frequency bands (chapter 8), the plan could guarantee only freedom from 3rd order passive intermods. But with a little ingenuity it was thought that 5th orders could also be avoided by specific allocation. The trouble with the higher orders was that, not only were they inherently wider spread in frequency, but that they were both individually wider when modulation was considered, and there were considerably more of them (see appendix J). The chances of producing a frequency assignment plan which avoided having to place receiver frequencies completely free of 11th order products when these conditions were taken into account was thought to be very small indeed. When other heavy constraints on the frequency assignment exercise were brought into the consideration then the chances were even smaller.

11.5.3 Computer Simulation

One of the points of high concern was that of the impact of other users of the site, or even other near by sites. Even one other emission at a site, or impinging on the site, could extend the spread of the intermod products considerably. This is shown diagramatically in appendix J.

A computer program was produced which attempted to simulate the production of intermods and display the result on the computer screen in the form of a spectrum analyser display. For a number of operator given frequencies the computer calculated all odd order intermodulation products from 5th to 11th, and displayed those which fell within an operator determined bandwidth as vertical lines. The hight of the line was determined by the order number, with lower orders being higher to illustrate the expected ordering of magnitudes. The operator could also choose to simulate modulated carriers which would cause the displayed line to have a width proportional to the chosen occupied bandwidth (say 8kHz for AM) and proportional to the order number. With this facility trial transmit assignments could be made and their impact on the receive bands assessed. There was the ability to magnify the frequency axis in order to examine the detailed structure if desired.

The processing time for the extensive calculation needed made the use of the facility limited to assessing only a few transmit channels (say 6). But it did give a good insight into the form of intermod location and the extent of their coverage. It further emphasised the severe impact of the site sharer - ie a site user radiating a signal outside the Directorate's frequency bands and over which we would have no frequency control. This is illustrated in figures 11.5.2 and 11.5.3. They are 'screen dumps' of the computer display. The first shows the intermod distribution for five transmitting frequencies, two in the lower transmit band of 152-153 MHz, and three in the upper band of 154-156 MHz, as they effect the receive bands from 144 to 148 MHz. If a horizontal line were to be drawn across the graph at the level of say the 7th order products then this would indicate in the clear portions the frequencies which would be unaffected by intermods of that level. As shown just over half the overall band is free of 11th order.

The second diagram (figure 11.5.3) shows the effect of adding just one transmission in the band between those used before and which would simulate the use of one of BT's paging frequencies on the site - as happens in practice. It will be seen that the picture is radically different, there is no part of the receiver band free of 9th order products. (Both diagrams assume modulation with an 8kHz occupied bandwith). Also shown on the bottom left corner of the diagram is the number of (carrier) products, which fall in the 4MHz receive band indicated, as a function of order number.

11.5.4 Outcome

To try to find solutions to the overall problem of intermods/frequency planning a "brain storming' meeting was held with a small number of parties who were both 'interested' and in a position to contribute ideas - however far-fetched. A large number of ideas were generated but most involved long term investigations to evaluate them and, with the previously quoted shortage of resources and time it was not possible to persue them. They are not reported on here for that reason, and because they were both frequently Directorate specific, and do not represent sole contribution by the author. Some suggestions were of immediate impact, like retaining two-tower sites wherever possible and using screened good quality aerials such as welded skeleton slots. The main outcome was however to put the solution to the problem squarely onto the shoulders of those involved in producing the frequency plan.

12 Frequency Assignment I (An Assessment of Possibilities)

12.1 Introduction

12.1.1 General

It is necessary to take the broad perspective previously given to the topic of operating frequencies and refine this into usable detailed frequency assignments which can be used on a national basis. Ideally this should have been part of the work in choosing new frequency bands since an examination of the details might show fundamental problems in the choice of bands. However, time scales dictated that the ideal could not be achieved. The aim of this chapter is not to carry out the detailed assignment but to first, using general arguments, see whether it is compatible with the desired constraints. Secondly to indicate, if necessary, where the constraints must be modified, and then give general guidance in the strategy to be adopted in such assignment, which will be carried out in chapter 13.

It is worth recapping at this stage on both the starting point, the old bands, and the allocations into which the services must be changed. These are shown in figure 12.1.1.

The end point of the assignment exercise is reached when all the radio communication services have been assigned a specific frequency (and operating parameters such as bandwidth, power, etc), and all their interactions have been taken into account. In the situtíon under study this means all police and fire service mobile channels, and all their associated fixed radio links. This is a total of some:-

250 police mobile channels
75 fire mobile channels
300 outgoing link channels) note some links are used in an onward
500 return link channels) repeater mode

These are deployed over England and Wales from

250 base station sites
100 control locations covering:-
 50 fire authorities
 43 police authorities

Base stations are frequently shared by both police and fire users (and others), and sometimes by two or more authorities.

12.1.2 Frequency Planning Aids.

Frequency planning and assignment is a continuing task within the Directorate of Telecommunications and is undertaken by one group (Frequency and site planning group - FSPG). They had recognised that the complete replanning of these bands, taking account of the constraints under which this would have to be done (12.1.3), was beyond their normal experience which was of dealing with either individual assignments on a piecemeal basis or occasionally negotiating new block allocations (for UHF links or microwave links for example). Consultants were therefore commissioned to define what would be needed by way of planning aids. By discussion with all involved parties the consultants refined and regularised the virtually instinctive modus operandi of FSPG onto a flow chart.

Very briefly, this consisted of making a trial assignment at a particular site and then testing this against a succession of constraints - such as those to be listed in section 12.1.3. Each test required access to a data bank of information and then, frequently, processing this into a usable form. Thus a computer would be required to store and access the data and perform the calculations, some of which could be lengthy. The author was involved in this formulation in a general consultancy role for the individual tactics to be used in each test.

The FSPG required that the computer and its program (provided by the consultants) be so configured that it would not make assignments itself. It was a requirement that the process be under the control and intervention of an operator. Thus the operator would initiate a step in the assignment processs and call on the computer to calculate whether the input quantities for that test passed the constraints involved. The operator also set the pass level of the constraint. If the test were successful, then the next step could be approached. If not successful then the operator had the choice of either relaxing the constraint or changing the inputs to the test. For some tests the constraint level could be an output of the test, eg 'test passed at levels 1-3, failed at level 4.' For others only one level could be assessed at a time. This applied particularly to those tests on high order intermodulation products.

Those tests could be particularly time consuming so when the author was given the task of 'second sourcing' a frequency plan (in a very short time scale), it seemed sensible to question the general philosophy of use of the planning aid.

Whilst the data bases were obviously of great value for their contents they were often in a form specifically designed to be of use for the one task of the computer - planning according to the FSPG philosophy. In addition at the start of the involvement, at least, they were incomplete. The greatest drawback to its use was the fact that apparantly many paths could be explored only to find that each could be a blind alley. Until the last country-wide assignment had been made, there was always the very real chance that this could be incompatible with one of the many constraints. If this happened, then it would be necessary to back-track, make changes, and try again. But there would be only the experience and feel of the operator available to guide in the choice of how far to back-track and what to change.

This seemed an unsatisfactory process therefore, at least for a person new to the art of frequency assignment. What was needed was some overall guiding strategy which could give a very high probability of ensuring a successful outcome. It is the purpose of this chapter to discuss such a strategy, and to develop it in chapter 13.

12.1.3 Constraints

The channel frequency assignments must be made against a number of constraints which are given below (in no particular order);

a) channel bandwidth to be 12.5 kHz for both mobile and links.

b) fixed links will use directional aerials wherever possible.

c) main transmissions/receptions can be assumed to be omnidirectional.

d) there will be many area coverage/quasi-synchronous channels.

e) the desirable minimum protection from cochannel interference for links is 18dB for 90% of time.

f) the desirable minimum protection from cochannel interference for mobile services is 18dB for 90% of time.

g) the Directorate's policy is to remove all fixed links from the VHF bands as soon as practical.

h) continuity of service must be maintained during the change over.

i) desirability to retain the existing link frequency assignments if possible.

j) allocations which give rise to 3rd order intermod products on other assignments at any one base station are to be avoided.

k) main receivers at base station are to be protected from intermodulation products of the colocated transmitters up to ideally at least the 11th order.

l) a major user of the 140/150 MHz bands already, for personal radio use, is the Metropolitan Police Force (London) and particular constraints are attached to this aspect.

m) some base stations are required to radiate transmission for more than one county.

n) an ability to accommodate future growth is highly desirable.

o) adjacent channels are not to be allocated to main transmissions within a county or from a site. Furthermore adjacent channels are to be avoided in adjacent counties.

p) desirability to avoid the use of the lower parts of the 152-153 MHz band near to Jodrell Bank and Cambridge to avoid possible interference to Radio Astronomy.

q) It is highly desirable that the plan be flexible to take account of changes which may be desired (eg growth and enhancements) and which may be imposed (eg as a result of national and international frequency clearance).

12.1.4 Comments on Constraints

Some of these constraints will be discussed in much greater detail in suceeding sections but a few of the others need further comment here:-

e) The figures for protection ratio are difficult to establish; the ones quoted here result from a combination of considerations covering:-

- what is thought to give acceptable communications quantity,

- extrapolations from current (and acceptable) re-use distances at 100 MHz,

- values which are covered by CCIR curves for expected levels.

f) Although the links are expected to be of better quality than the mobile segment, at least for the latter (at the extremes of range) the fact that their variability is much less and the small amount of capture effect which is left at 12.5 kHz channelling, means that the protection ratio in terms of RF carrier ratios is comparable to that of the mobile segment.

g) It is difficult to assess how many of the links would be converted to other bands (expected to be at or near micorwave frequencies) during the change-over period. It depends on the availability/suitability of new bands, and the manpower and financial resources which can be devoted to the task. Prudence suggests that this factor be used only as a bonus to cope with those links which otherwise would cause problems (eg onward repeaters) and to provide the possibility of more mobile channels in the future (ie growth) when the link conversion is completed.

l) The London scene is different to the rest of the country. They currently use a large number of personal radios in this band, working to many local base stations - often located on police stations and thus not of the same nature as hill top sites. The radiated powers are low to cover the short range required, and allow reuse of channels within London. There is a significant problem, however, in that they transmit in the 140 MHz band from the base stations and thus break the convention of 'base station high' (in frequency). It has been agreed that they will conform to the general policy for the WARC frequency change, but this introduces particular problems during frequency conversion when some mobiles will be operating 'base high' and others 'base low'.

m) The sharing of sites between adjacent counties leads to a chaining of constraints; this is illustrated in figure 12.1.2 (the details of the constraints are discussed in the next section 12.2). Consider one of these sites, then all the radiations at the site will have to conform to a particular frequency pattern (as will be seen in detail later in this chapter). Since other base stations in the county will most probably carry some at least of the channels assigned to the first one, then elements of their assignment pattern have been determined. Similarly, since the sharing ones are part of the initial channel pattern, they carry the seed of another pattern into the base stations in their county. If one or more of these other base stations again shares with yet another county then the chain is continued. Thus the frequency assignment patterns of the sites within one county have to inter-mesh with not only their co-county ones but also sharing counties. This effect has been termed cogging from the analogy with meshing gear wheels.

12.2 3rd Order Intermodulation considerations.

12.2.1 Main Transmissions

Section 11.1 has described how intermodulation products can arise within the mobile receiver. If a mobile is close enough to a base station, then when two or more base station transmitters are activated they will give rise to intermodulation products, if a product were to fall onto the frequency to which the receiver was tuned then its effect would depend on firstly whether that channel was also being radiated and secondly the intermod level.

If the wanted channel is radiated then the intermod level can confidently be expected to be well below it in level and have negligable effect. If, however, the wanted channel is not radiated then again the intermod will have no effect until it is strong enough to open the receiver's mute. When this happens the mobile user will be subjected to the demodulated intermod product which may be at least partially intelligible or, perhaps more disturbing, recognisable as two or more speech messages which cannot be deciphered.

It is, therefore, highly desirable to avoid channel assignments at base stations which give rise to such intermodulation frequencies, otherwise there is a zone (approximately a circle) around each base station within which the mobile would experience highly undesirable interference. The size of the zone is dependent on the transmitter powers, the performance of the mobile receiver, and the order of the intermod considered. The zone might have radius of some half mile for 3rd order products and it is usual to ensure that assignments are free of this order and perhaps 5th order as well.

The assignment pattern needed to avoid 3rd, and to some extent 5th, order products has been studied (Edwards, Durkin and Green 1969) and they, and others, (Panell 1979, Quarell 1983) have produced lists of suitable channel allocations. Edwards et al showed that the criterion for third order avoidance was to ensure that no two channels had the same difference in frequency as any other two. The consequence of this is that the frequency range needed to cover the assignments at one site (sometimes called the switching range) increases rapidly with the number of channels required. This is shown in fig 12.2.1 using values from Edwards et al and Quarell (Panell allows the use of adjacent channels which is deemed not advisable at least for the emergency services). There is no proof that the values shown are in fact the minimum ranges, since they have been derived by setting a computer to search for sequences which met the difference criterion. The number of operations to find sequences close to the minimum also increases rapidly as the number of channels is increased. This necessitates many hours of computer time, even for a moderate sized mini computer (eg PDP 11/23) on this sole task. Smaller ranges might be found given a more diligent search, but the values given are thought to be at least close to the minimum and it is difficult to justify more effort for the possible refinement of the figures. It is worth noting that Quarell has reported that the suitable sequences can be found more easily (ie more quickly) when relaxation is given to being close to the minimum range. A consequence of this is that, whilst there may be only one or a few sequences of minimum range, the number of sequences of a slightly larger range is very much higher and there are indications that the number escalates rapidly with distance from the minimum. Examples of minimum range sequences are given in Table 12.2.1.

The consequences of taking into account 3rd order free assignments are therefore a commitment to a spread in frequency of the assignments at any base station. By way of example, 8 channels which occupy 100 kHz of actual spectrum would require 500 kHz to be available to avoid 3rd order products.

Table 12.2.1

Number of channels	Channel Assignments
3	1,3,6
4	1,3,6,10
5	1,3,7,10,15
6	1,4,8,13,19,21
7	1,3,11,17,22,26,29
8	1,4,9,15,19,31,38,40
13	1,3,6,10,23,31,42,54,69,87,111,121,127

An empirical expression was found which was a good to the points shown in figure 12.2.1 this was

$$Z = 11.56 \times 10^{.0837n} - 15$$

where Z in the number of channels required (the switching range), and n the number of channels to be assigned. This curve enabled extrapulation to be made to values higher than 14 wanted channels.

12.2.2 Link transmissions

It is normal in mobile radio design to consider the 3rd order implications of the main transmissions only, but it is more realistic to take into account all transmissions which could reach the non linear stages of a mobile receiver. Frequencies away from those of the band of operation of the receiver will be attenuated by both the input stages by virtue of the filtering action resulting from a need to tune and match them, and the resonance of the receiving aerial. On this basis emissions other than those of the main transmissions are usually ignored. Although troublesome situations can arise in the vicinity of powerful transmitters, for example those of sound and television broadcasting.

In the situation under consideration here, however, the fixed links will share the same frequency band and their effect must be considered. The transmitters are of the same order of power as the main transmitters but they are connected to directional aerials. It might be thought therefore that any effect of the links would be felt in the main beam of the aerial. For the six element Yagi aerials envisaged the main beam is some 40° wide, so that certainly within this beam, strong signals can be expected due to the aerial gain. The vertical directivity is unlikely to have any significant effect, but the use of horizontal polarization could be of some benefit in reducing the level received by the mobile's vertical aerial to that of an omnidirectional aerial. Outside the main beam, the levels of side lobes cannot be relied on to be more then some 10dB below the main beam. No polarization effects can be counted on here, so that again, the received levels will be similar to that of an omni aerial. In consequence, a reasonable assumption is that the mobile will experience a level into its receiver from the link transmitters similar to that from the main transmissions.

12.2.3 Combined Effects

Thus both main and link frequencies must be considered as potential 3rd order interference sorces and, therefore, their assignments need to be linked into a common 3rd order free sequence.

For the site with 8 main transmissions, there will be 8 return link frequencies (on the most reasonable assumption that there will be as many main receive channels as transmit channels at the base station) and in fact the fire service return links must also be taken into acount for which there could be probably two at such a well used site. Thus the requirement jumps from 8 channels 3rd order intermod free to 17 or 18!

Such a site might be thought to be rare, but figure 12.2.2 shows a histogram of number of sites against number of 150 MHz channels. It will be seen that there are a significant number with a requirement for 14 or more channels, and 14 is the highest number for which we have an assignment sequence and even this requires a range of 157 channels, ie just under 2.0 MHz at 12.5 kHz spacing.

12.2.4 Consequences

It will be remembered that the links already exist in these bands and that it is desirable to retain their allocation. However, the chances of these allocations matching a desired sequence are so remote as to be negligible. [The existing allocations were determined in an ad hoc manner with no requirement to meet the 3rd order free criterion]. Thus the links must be reallocated if the 3rd order considerations are to be met.

This poses practical problems for implementation. The current allocations have only been determined after confirmation of the suitability of an assignment by a practical trial to see if there were any interactions to or from cochannel or adjacent channels links (the link frequencies are reused many times). This takes time but there will not be such time available during the WARC change-over.

12.3 High Order Intermodulation Considerations.

12.3.1 Initial View

Section 11.5, the implications of passive intermodulation products, showed that the base station receivers would be vulnerable to intermodulation products up to at least the 11th order. It also demonstrated how the receive bands were vulnerable to an allocation of transmitter frequencies which were spread widely.

This is the situation in which we now find ourselves. The 2 MHz occupancy of the transmitters means that, even if they are located within the contiguous 2 MHz of the transmit band, the 11th order products will spread to 10 MHz below the bottom edge of that band. The result of this is that the upper receive band will be completely blighted with both 11th and 9th order products as shown in fig 12.3.1. Considering the number of transmissions, and in consequence the number of intermods, the chances of finding an unaffected receive channel in this upper band is very small, and if the spreading due to modulation is considered then the chances become even smaller

Nevertheless, the lower receive band remains free of troublesome intermodulation products and the requirement for so many channels to be radiated will be met only infrequently, so these can in principle be deemed to fall into the category of having to be provided by other means (eg microwave links).

12.3.2 Further Considerations

There are two other factors however which counter the amelioration just discussed.

Firstly, the chaining of counties will make the search for compatible sequences for all these counties difficult - bearing in mind the need to avoid adjacent channels in adjacent counties. This is likely to mean that the lower 150 MHz band (152-153 MHz) will have to be brought into use with the consequences to be described in section 12.3.3.

The second factor is that of the practical implication of changing the link frequencies which section 12.2.4 has shown to be necessary. Whether conducted during the actual WARC changeover (originally the main transmissions only were to be changed) or before it, the operation will be very difficult since the present links are spread over some 1.8 MHz of spectrum in the 154-156 MHz band. This is due to the fact that the use of the band was not planned in its entirety from the outset; each link appears to have been a special consideration which had to fit in with other links and, at that time, other services. The original allocations were for 25 kHz channels (reduction to 12.5 kHz channeling was made policy at about the start of the WARC considerations) and as the band became more crowded each new provisional assignment was tried gingerly for compatibility with existing services and only confirmed after a successful trial.

There is therefore, a need for elbow-room when making the change and the obvious place to find this is the 152-153 MHz band which is currently unoccupied by other transmissions (having been reallocated from the radio astronomy bands). On both counts then, the bottom, isolated, 1 MHz of transmit band must be used. A double move into this band and out again to 154-156 MHz is disadvantageous.

This now very much complicates the use of the receive bands when high order intermodulation products are concerned. Just one assignment at the top of this 152-153 MHz band coupled with the previous case of other assignments, which will spread anywhere in the upper 2 MHz band, result in the situation shown in figure 12.3.2 which shows the receive bands completely blighted with products from 9th to 5th order. The magnitudes of the lower order products are likely to be particularly damaging.

This is perhaps a rather unrealistic view in that all the links would have to move to the lower 1 MHz band leaving the main transmissions in the upper portion. The total spread of frequencies in the two bands cannot be less than that given by the 3rd order free sequence for the total number of transmissions. Since these transmissions are now in two parts, careful consideration must be given to this separation in order that the allocation criterion is not broken.

The spread of intermodulation products into the receive bands now becomes more difficult to predict and depends on the actual location of the sub bands and their width. An example can be given for the case previously considered under the most favourable circumstances. This is shown in figure 12.3.3, where 7 channels have been allocated to the link (lower portion) band and this and the upper band for main transmissions are each deemed to occupy 1 MHz.

The figure shows only one location for the upper 1 MHz of the transmit band (from 154-155 MHz), but other locations are possible. This is shown in Figure 12.3.4 which follows the form of appendix W. The figure indicates the position of the intermodulation products as a function of the location of the upper 1 MHz. Since it might be considered that a too restrictive case had been considered, then figures 12.3.5 and 12.3.6 show the situation for 0.5 MHz of main and link bandwidth, for the location of the 0.5 MHz in the lower band. The situation can be seen to be far from favourable.

12.3.3 Interim Conclusions

It has been shown that the fixed links cannot remain on their existing allocations if the 3rd order free constraint is to be met. Furthermore, there is no certainty that a diligent search for suitable transmit allocations and receive allocations would yield an acceptable national plan with a minimum of fixed links having to be rehoused elsewhere in frequency.

An alternative strategy is therefore highly desirable.

12.4 Reconsideration of 3rd Order Requirements

12.4.1 General

The present difficulty is generated by two conflicting requirements. To avoid 3rd order products the transmit frequencies need to conform to a distribution pattern and this dictates that, for moderate numbers of channels, they are dispersed in frequency terms and cover a wide bandwidth. Transmissions which cover a wide bandwidth give rise to a high order intermod products which fall onto the receiver frequency bands and occupy them to a degree which is unacceptable, both in terms of bandwidth and levels. The practical need to relocate existing fixed link frequencies into the lower (152-153 MHz) transmit band, in order to avoid clashes during any changeover and meet the distribution pattern, turns a very poor situation into a virtually impossible one.

To overcome this it is necessary to break either the tie between these factors or one or more of the constraints themselves. The linking seems firmly established, and no way can be seen of tolerating or eliminating the high order intermod levels. Thus attention is directed to the 3rd order factor.

12.4.2 Link Aspects

The strategy which has been adopted so far in assigning the transmit channels has been that of avoiding any third order products falling on a non participating channel (ie one which is allocated at that locality but not involved in generating that product and might therefore not be transmitting). An alternative strategy can be proposed as one which seeks to maximise the number of 3rd order products falling on any nominated channel - presumably also one which is not a transmit channel. By this means fewer channel will be affected by intermods and hence the sequence can be shorter.

To accomplish this then, a number of difference frequencies (ie the difference in frequency between any two transmit assignments) must be the same. A sequence of this form is shown diagrammatically at the top of figure 12.4.1. The spacing between A and B is the same as between C and D, and also between E and F; the separation of these pairs is not important at this stage. The resulting 3rd order intermod spectrum is shown with the contributing channels indicated; other 3rd order products outside this sequence are not shown.

This has achieved the desired end of ensuring that the intermods tend to accumulate on some channels and indicates that the frequency span (F-A in Hz) will be less than that of a corresponding 6 channel 3rd order free sequence. [A spacing of 2 units (B-A being one unit) between B and C, and 3 units between D and E ensures no further interactions and gives an overall length of 8 units compared to a normal 3rd order free sequence of minimum length 21 units.]

The stated aim of avoiding other transmit channels has not been achieved however. This of course follows from the fact that to avoid such occurances the spacings must all be different as has been stated before. In fact not only do the intermods occur on those channels but they are in mulitple numbers there also.

But this is in fact of no consequence if the interfered with channel are those assigned to link frequencies!

The mobile receiver, whose non-linerarity gave rise to the requirement in the first place, will never be tuned to the link frequencies and therefore its mute will never be opened by the generation of intermods on the link frequencies. The generation of such intermods in the distant link receiver which is tuned to these frequencies is also of no consequence, since a) the link transmitters are always on, and b) more importantly the received levels are so low as to make the intermod levels entirely negligible.

Thus the link frequencies can apparently operate on highly disadvantageous spacings so far as 3rd order products are concerned. For example they could be on adjacent channels. This would be acceptable for the associated link receiver since it would receive all the link transmissions at the same (low) level and they are specified to tolerate an adjacent channel upto 70 dB higher in level.

The mobile receiver is however not yet completely clear of the possibility of receiving signals which could generate intermods. If the spacing between any two link frequencies is the same as that of any two main transmissions, then the mobile receiver could experience mute opening on a non transmitting channel. This can be avoided by making the minimum separation of the mobile sequence (which must still be a 3rd order free one) just greater than the maximum separation of the link frequencies. For example for 8 link channels, then their maximum carrier difference is 7 channel widths (assuming tightest packing) so that the minimum main sequence separation would be 8 channels. But this is virtually the only restriction, so it is only necessary to search for a sequence for the, say 7, main channels which complies with this and otherwise treats the 7 channels as an entity; rather than trying to find a 15 (8 + 7) channel intermod free sequence. The 'virtual' comment here covers the one other restriction which is that the separation between the link transmit frequency band and the main transmit frequency band is at least the width of the larger of the two; this avoids the intermod products caused by one band itself which fall outside that band overlapping into the other band. This also takes account of the interaction between two channels in one band and one in the second.

This separation requirement can be fairly readily met even for the case of 8 links and 7 mains if it were desired to operate entirely within the upper 2MHz transmit band. But the need for the link frequencies to be located in the lower 1MHz band has already been discussed (12.3.2 and 12.4.1). This offers the opportunity of locating all the link base station transmissions in the lower transmit band and all the main tramissions in the upper 2MHz band. Providing that the link transmissions do not range over more than 1MHz (this will not happen if they are bunched) then there is complete freedom to locate the links anywhere in its band, and the mains similarily. Thus 8 main transmission can be easily accommodated on this basis.

12.4.3 Consolidation

The picture for allocations now appears, for the high usage site, as some 100 kHz of adjacent channel assignments for fixed links in the lower transmit band and some 1 MHz of bandwidth in the upper 2 MHz, assigned to the main transmission using a 3rd order free sequence.

The high order intermold phenomenon is still a problem however. The spread of intermodulation products in the receiver bands, whilst reduced compared to that before compacting the link bandspread, is still extensive to the degree that no reliance can be placed on finding sufficient free spectrum to be sure of being able to make a receive band frequency assignment. The intermodulation aspects are shown in figures 12.4.1, 12.4.2, 12.4.3. 12.4.4 for various locations of the 100 kHz in the link band and figure 12.4.5 for an illustrative 0.5 MHz main transmit bandwidth.

12.5 Reconsideration of Mobile Aspects

12.5.1 The Choice

Although the range of frequencies allocated to links at any one base station has been reduced, by showing that the third order constraints need not apply in their case, the spread of main transmissions due to this cause still poses a great problem for the usability of the receive bands due to high order intermodulation products. The only way that the spread of these main bands can be reduced is to break the constraint that the transmit sequence be 3rd order free. If this is done then the mobile will experience interference due to this cause.

The choice then is; should the system be configured such that the base station receivers are free of high order intermodulation products, or should the mobile be free of 3rd order products. The factors to be considered will now be discussed.

12.5.2 Impact on the Base Station

The base station could experience high order (upto say 11th order) products whose levels could exceed the desired protection ratio (30 dB) for the smallest signal it is designed to receive (from a mobile at greatest range). The levels could well be much higher than this which would give four modes of operation.

i) When receiving very distant mobiles - signals unreadable.

ii) When receiving moderate distance mobiles - signal suffer some interference.

iii) When receiving from average to near mobiles - no significant interference

iv) when no mobiles are transmitting on that channel then the mute could well be lifted giving rise to audible interference or even partially intelligible transmissions.

A time factor comes into play here on two counts. Firstly the interference can only be present when two or more transmitters are active [the outgoing channel that is paired with the incoming one under consideration can be one of the contributing transmissions since it will radiate pip-tone during reception]. It might be considered possible under such circumstances to seek an interference free channel but that means that all mobiles on that channel must change (entirely a time penalty) and there could well be interference on the new channel within a short period due to the dynamic situation of random channel activity. In addition to the difficulty of such a change during the occasions when the system is most vulnerable to any interference (heavy channel usage due to a major incident) there will be no spare channels to change to, and interference will be at a maximum since all transmitters are likely to be active. The second time dependent factor is that inherent in the high order intermodulation situation which is that the levels are not steady - they vary with time due, it is thought, to variations in tower conditions caused by precipitation, temperature and wind. It will be remembered, however, that the aim is for a grade of service which is high, at least in terms of chances of achieving an intelligable transaction. There are already a number of time/location factors which have eaten into any margin here, so that little is available for assigning to the intermod variability. Thus the higher recorded levels are of importance.

Examining the four modes above shows that obviously mode 3 poses no problems. Whilst this covers the short to average range, if the mobiles are evenly distributed by area then the more extremes of range cover more mobiles, thus most will be in modes 1 and 2.

Mode 4 is in many ways the most damaging of them all. It means that one person, the controller of that channel (and by implication of course each of the other controllers), although ideally situated in a low acoustic noise environment and serviced by high quality equipment, will be subjected to sporadic periods of interference. This will make his task very waring, leading to fatigue and probably an annoyance with the system. Thus his working efficiency could be low, and he is the central point or focus for that channel, also he can exercise no control over the onset of the interference.

12.5.3 Impact on the Mobile

It will be recalled that the mobile experiences 3rd order intermodulation interference when it receives two or three transmissions at a high level - that is when it is close to the base station.

If there is a transmission at the time on the channel to which it is tuned then the interference will be negligible. It only becomes obtrusive when there is no wanted transmission from the base station and the mute is opened.

Whilst this will be as disturbing as for the situation of the controller, considered in the previous section, it will occur only for the very small fraction of the total area covered by that base station. In addition, due to most base stations being located on remote hill tops, the need for, or chances of, the mobile operating in this intermod region are small. Even if experienced any one mobile (or the totality of mobiles) is unlikely to remain in the area for a significant period of time.

Moreover the mobile has a simple remedy for the interference - this is to increase the setting of the mute level. On many of the sets of mobile equipment in present use the mute setting is presented as an operator control variable and can be readily altered.

If the problem is as small as is now presented, then some explanation is in order as to why it has not been dealt with as such in the past.

The explanation must lie in the fact that the bulk of mobile radio users, other than the emergency services, have different operating conditions. Few will have their base station located on a remote hill top - firstly with the transmitter powers permitted they may well not achieve coverage of their local areas (if they did then the regulating authorities may deem the coverage to be too large for good spectral efficiency), and secondly there are the cost/feasibility/ licensing aspects of their fixed link. Such users are often accommodated on sites nearer to their controls (work premises) which implies more in urban areas and more users per site. Thus they are more likley to have mobiles close to the base station and greater likelihood of undesirable assignments. The allocation of such users on an 3rd order intermod free basis (implying that only a few of the available site channels are useable) is not so spectrally disadvantageous as may at first sight appear. Those users will not have area coverage requirements so that the unused channels can be assigned to another base station within a short range of the first.

In the case of the emergency services, there was the desire to give as good communications as possible, there were no high order intermod problems when operating in the old bands, and the spectrum was available to grant suitable allocations. There is also the very real danger that the user will increase the mute setting when experiencing interference and neglect to restore it when outside the intermod zone, and thus miss messages. This factor must be considered in the present situation.

The considerations can be put under three headings:

i) education/discipline
ii) penalty of change
iii) automatic mute setting

i) It should be possible to educate the user to restore the mute when outside the intermod region and back this with the stick of disciplinary action if not complied with.

ii) Since the new bands of operation are smaller than the old ones, and unfavourably spaced, the user must be prepared to accept that the service is unlikely to be better - this situation is one where it will be worse.

iii) It should not be impossible to devive a means of coping with the variable mute setting. Tone encoded mutes cold be advocated and would provide some help but they impose an overall penalty by having to reduce signal modulation in order to admit the tone, and the tone could be overidden for moderately high levels of interference. A better solution would appear to be that based on an out of (audio) band noise sensor (possibly combined with a tone mute). A very simple solution would be to arrange for the mute to be restored after a given period of time. All these pose cost questions (which return us to the penalty consideration) but provide a rich field for further investigation.

It should be pointed out that raising the mute in this region will of course, in itself, not degrade any wanted messages since they will be conveyed by RF signals which are themselves strong.

12.5.4 Other 3rd Order Factors

The discussions on 3rd order intermodulation factors has concentrated on the situation where the intermod product is generated in the mobile receiver due to its inherent non linearity. But 3rd order products will be radiated by the base station by means of the passive mechanisms described in chapter 11. In this situation those products generated by the site non linearities will be only of any significance very close to the base station indeed.

Higher level 3rd order products are however likely to be radiated by the mechanisms of transmitter to transmitter output stage coupling via the transmit aerial(s), as described in section 10.5. Such levels may be only 30-40 dB below the wanted carrier and could give rise to mute opening, in the mobile receiver over a much greater range than that caused by the mobile receiver's non linearity. This effect can be overcome by either using inherently more linear output stages than is normal in mobile radio (use of amplitude modulation could well be beneficial here since it inherently requires a greater linearity of output stages than does FM), or by adopting techniques to block the intermods.

Improved output stage linearity is obtained by either the use of the multichannel transmitter or, for single channel transmitter, by the use of output feed back techniques. all that is required to achieve this is to call for the desired performance in the purchasing specification (and pay the consequential costs). This can be contemplated for the case of the main transmitters which will be new items but not for the link transmitters which will be (largely) the existing ones.

For the link transmitters it will therefore be necessary to adopt intermod blocking techniques. Narrow band pass filters would seem to have no part to play here. The requirement is to cater for even adjacent channel operation. But isolators will work in these circumstances, although they can only be counted on to reduce the product level by one dB for every dB of isolation.

It is however possible, at least in principle, to achieve sufficient reduction in intermod level by putting in enough isolation (whether by isolators, aerial separation, or combining techniques). Thus the site non linearity problem turns into one of (extra) cost for the provision of good link transmitter aerial coupling design.

12.5.5 Outcome of Reconsiderations

Consideration has been given to a number of factors which affect the choice, given at the start of section 12.5, as to whether the base station, and hence the controller, should be expected to tolerate interference, or whether the mobile was in a better situation to shoulder such impairments.

The arguments point clearly to the fact that not only should the mobile segment be better able to accommodate the interference to which it would be subjected, but that there should be high hopes of finding ways of lessening the impact of such interference - to the extent of making it negligible.

A decision can therefore be made to break the constraint of having to assign the outgoing channels as a third order free sequence. On this basis the mobile channels can be regularily spaced (ie every th channel can be assigned) even to the extent of radiating on adjacent channels. This would give the ultimate in narrow band operation, and it is necessary to see whether this now admits sufficient freedom of high order intermods in the receive bands to be able to contemplate assigning receive channels.

Fig 12.5.1 shows the high order intermod situation in the receive bands on the assumption of 8 link channels and 8 main channels with each of them being assinged on an adjacent channel basis within their own bands. That is 100 kHz of link bandwidth and 100 kHz of main transmitter bandwidth; the former in the lower transmit band and the latter in the upper transmit band. Figures 12.5.2 and 12.5.3 are similar with different locations of the link band within 152-153 MHz.

The indications are that sufficient bandwidth will be available in the receive band.

12.6 Conclusions

This chapter has sought to answer the question: 'Is it possible to show whether, taking account of certain constraints, it was sensible to start a frequency assignment exercise for the base station transmit (main and link) and receive (main and link) 140/150 MHz bands'?

The initial conclusion was that there was a very low probability of the assignment exercise achieving its goals within the constraints. These constraints were then examined to see which could be broken, and that of 3rd order free transmit channel allocation was shown to be discardable. Under these conditions the assignment exercise could be contemplated. If successfully completed, then individual sites which may have most unfavourable allocation or be most vulnerable to the breaking of the constraint (base stations in populated areas are a case in point), could be considered for realloaction on the assumption that the initial plan will be flexible enough (by virtue of the narrow band operation) to accommodate this.

13. Frequency Assignment II, A Plan

13.1 Recapitulation

13.1.1 Introduction

The previous chapter examined strategies for frequency assignment and found one which showed promise of yielding a result. To get even to the stage of being able to start the detailed assignments rquired that some of the constraints, cited at the start of that chapter, be broken. The chapter concentrated on the topic of assigning base station transmitters. The assignment aspects of the associated base station receivers was included only on the grounds that it provided a constraint to the transmitters via the high order intermodulation phenomenon.

The end point of that chaper was the defining strategy of minimising the spread of link channel assignments by assigning all link frequencies on closely spaced channels (to the extent of adjacent channels), and locating them in the lower transmit band (152-153 MHz). Similarly the main transmissions at a site were to be as closely bunched as possible - again to the extent of using adjacent channels since the 3rd order free constraint had been broken. On this basis sufficient of the receive bands should be free of the high order intermods to allow assignment of receive channels against their constraints.

13.1.2 An Illustrative County

A skeleton picture of the disposition of base stations, controls, fixed and mobile links was given in chapter 3. Some flesh can usefully be put on this diagram in order to show the essential elements of a county structure. This is shown in fig 13.1.1.

The controls are shown as triangles with the letters P and F depicting police and fire respectively. The base stations are denoted by squares and are lettered A B C D. The lines connecting controls and base stations represent the fixed link paths (assumed here to be both go and return) and the number alongside each indicates the number of channels carried on that path. In general the base stations are connected only of onward linking from B to C so that B to their own police and/or fire control, but the diagram shows one case becomes a master station. Also shown are links to A and D from other counties. Alongside each base station the numbers circled represent the 150 MHz mobile channels radiated, they are the police force's local nomenclature for channel numbers so that every force will start its sequence from 1. Fire service mobile channels are not shown since they will be at 70/80 MHz and are not critical for current planning purposes. The fire service link requirement is shown however, since it will use the 140/150 MHz spectrum. It should be noted that site D is called on to radiate 16 frequencies in the 150 MHz band.

The diagram carries all the essential elements to form a building block for the national requirement and will be used for illustration throughout this chapter.

13.1.3 Onward Links

This chapter will need to consider the details of assigning specific channels (frequencies) to both the base station transmitter, links and mains, and base station receivers, links and mains. In doing so it will automatically take account of the transmitters and receivers situated at the control locations since these form the far end of the base station links. There are however exceptions to the general rule that links are used only for control to base stations; in a number of situations they are used for 'onward linking' or in a 'repeat mode'. Thus a signal to or from control may be sent to a base station on a fixed link, perhaps used there, but sent on to another base station which cannot access control directly via a fixed link. This is shown in fig 13.1.1 for the link from B to C. The first base station is then often termed a master station, and there can be several 'hops' of onward linking. This requirement poses headaches for operation in the present bands because it means that at one or more of the sites either a link transmitter needs to operate in the link receive band or a receiver to operate in the transmit band. Either way a receiver needs heavy protective filtering to prevent interference from the transmitters, and places restriction on the frequency assignments.

Because operation in the new bands will be more difficult than that in the old, and the task of frequency assignment more onerous, a decision can be made that these onward links will be ignored at least until a plan has begun to take shape or even finalised. Then their impact will be reviewed, and the choice made of trying to fit them in or assigning them elsewhere.

13.2 The Geographical View

13.2.1 The Counties

The disposition of the police authorities in England and Wales, which are served by the Directorate for planning purposes, is shown in fig 13.2.1. Some authorities cover more than one county, examples are Devon and Cornwall, Avon and Somerset, Thames Valley .... Such forces tend to have requirements for large numbers of channels and a large number of base stations. This is often a consequence of their history of formation (a coming together of formerly separate forces in the Local Government Reorganisation of 1974) as much as a need to cover the population and topography. Although not strictly counties, therefore they will often be referred to as such, rather than an authority, for convenicence.

13.2.2 Security

The disposition of base stations, the details of channels they handle, and the link paths are of course available for planning pruposes, but are covered by a security classification. It is not possible therefore to reproduce them here without classifying the whole document. It is not necessary, however, to use these in order to illustrate the principles evolved for frequency planning. Artificial structures can be used - such as that of the illustrative county of section 13.1.2.

13.2.3 Controls and Base Stations - The Links

Figure 13.2.2 shows a idealised form of England and Wales. The headquarters for each of the idealised counties is shown as a filled circle and the base stations as open circles. The fixed links are represented by the solid lines joining headquarters and base stations, with again the number alongside representing the number of channels being passed over that path - at present on separate RF channels. The idealised map cannot show (at this scale) the number of fixed links - at the start of the planning exercise, these were numbered at 500, a figure which was to grow as the records were updated.

13.2.4 Chaining

A number of the counties share one or more base station sites. This was shown at A in fig 13.2.1 and is indicated on fig 13.2.2 by links from controls in different counties drawn to a base station. This means that these two counties are tied to one another for certain frequency planning aspects. Originally this would have been the common 'seed' to the 3rd order free sequence as described in sections 12.1.4 and 12.3.2. Since the third order free constraint has been dropped, this no longer applies in that form, but the desire now is to minimise the spread of channel assignments at hill top sites so that a tie between the two counties still exists. The second county of the tie may well share with another county, thus leading to a chain of tied assignments.

There are in reality a number of separate chains some of only two counties, some of four or so, and one very long and involved one of many counties. The larger ones have several branches and sometimes these branches, or the whole chain, can form a loop. Such patterns are shown again in idealised form in figure 13.2.3 where the lines connect the headquarters sites of the counties involved.

13.3 Assigning the Transmit links

13.3.1 The Requirements

It will be recalled that the desire is to locate all the base station transmit links in the lower transmit band of 152-153 MHz. There is also the desire to allocate up to 0.3 MHz of this band for the Metropolitan Police requirement. This 0.3 MHz can of course be reused away from the London Area but its manner of usage will not be for links, and it could be at fairly high power. Therefore it was deemed prudent to consider initially that it would not be available elsewhere in the country - it could be kept in reserve in case of problems. The aim then, is to restrict the link overall bandwidth to some 0.7 MHz - 56 to 60 channels say, and to try to restrict the assignment at difficult base station to as narrow a bandwidth as possible.

The first question to be settled is; can the links be contained in this bandwidth, it implies a reuse factor of 500/56 ie 9 to 10 times for the most part. It has been said previously that the present links had been planned one at a time and tested, with the whole number taking many years. A technique was therefore needed for quickly planning such links.

In fact two techniques were found and ultimately used in conjuction.

13.3.2 The Keyhole

Since all the link paths were already in use, but at an incorrect frequency, then there would be no problems with determining whether the desired path and received signal strength could be achieved. Problems would come from cochannel interference if link paths on the same channel were too close - but how to determine this distance?

The situation is depicted in figure 13.3.1. This represents a scaled map with a link on one frequency going from A to B. A second link on this frequency can be implemented from C to D if the level of A's signal at D is below its wanted signal from C by the protection ratio Γ. For the narrowband FM links to be planned, this is taken as some 20dB. [The exact figure is ideally a trade-off between the interference due to this cause and the interference due to all others but the trade-off cannot be performed at this stage so a nominal figure is taken. Indeed the trade-off would involve a very considerable amount of calculation and is therefore not worthwhile.] There is a similar consideration for the signals received at B due to the wanted from A and the interferer from C.

The path lengths are d_a and d_c from transmitters A and C respectively to other wanted receivers, and the interfering path under study from A to D has a length d_i.

The conventional approach is to calculate the transmitter powers at A and C to give the desired receive level at B and D using CCIR curves (CCIR 1978b) which can allow for obstructed paths and anomolous propagation. Then, using the CCIR values, the level of A at D can be calculated, taking account in all cases of the transmitter and receiver aerial gains in the relevant directions. This would be a very lengthy process for each of the 10 or so links on each of the 56-60 channels. If the calculations showed an unacceptable level then a more distant link would have to be chosen and the recalculation made. On the other hand, if the interference level were acceptable, then perhaps closer packing could be achieved for a nearer link path - another calculation.

What was needed was a way of quickly predicting which link path was the nearest that could reuse a frequency to that assigned to A-B. Or, since the closest might not give the best overall result when many cochannel links have to be taken into account, what technique could be used for determining the maximum reuse of a link frequency on a geographical pattern of pre-determined link paths. The technique developed derives from a simplified consideration of the lengthy calculations described.

The link transmitter power is determined by the formula:-

$$P_t = \frac{P_r\ F(d)}{K\ G_1\ G_2\ \lambda^2}$$

where P_t and P_r are the transmitter and receiver power respectively,

F(d) is the propagation law as a function of distance (d^2 for inverse square)

K, a proportionality constant, G_1, G_2 the gains of the transmitter and receiver aerials respectively, and the operating wavelength λ.

G_1 and G_2 are more properly a function of the angle at which the gain is measured relative to the angle of maximum gain, ie $G_n(\alpha)$, such that

$$G_n(0) = G_n$$

The level of the interferer received at D from A is therefore:-

$$P_i = \frac{P_{ta}\, K\, G_a(\theta)\, G_d(\phi)\, \lambda^2}{F(d_i)}$$

where the suffixes denote the particular location for that parameter. The aim is to find di such that the protection ratio criterion is just met, ie:-

$$P_i = \frac{P_{rd}}{\Gamma}$$

Now the links are designed to achieve a particular level of signal into the wanted link receiver. This is 30 μVemf or -107dBW. This is both necessary, to obtain the signal quality desired, and sufficient, since any larger value cannot be sensibly used and will imply higher than necessary transmitter power; with the consequence of interference over a needlessly wide area. Therefore:-

$$P_{rd} = P_{rb} \text{ and hence } \Gamma = \frac{P_{rb}}{P_{ri}}$$

thus

$$\Gamma = \frac{P_{ta}\, K\, G_a\, G_b\, \lambda^2}{P_{ta}\, K\, G_a(\theta)\, G_b(\phi)\, \lambda^2} \cdot \frac{F(d_i)}{F(d_a)}$$

$$= \frac{G_a\, G_b}{G_a(\theta)\, G_b(\phi)} \cdot \frac{F(d_i)}{F(d_a)}$$

In practice the aerials are six element yagis and it can be assumed that they are similar in performance, so that the suffix can be dropped. Thus the interfering range can be determined from

$$d_i = F^{-1}\left[\frac{G(\theta)\, G(\phi)}{G(0)\, G(0)}\, \Gamma\, F_{da}\right]$$

Now $\frac{G(\alpha)}{G(0)}$ is of course the polar response of the aerial referred to the maximum gain. A typical pattern for a six element yagi is shown in figure 13.3.2.

The wanted link path is most usually clear of obstructions, and virtually line of sight propagation is achieved between the two well elevated aerials, thus the propogation law for the wanted path can be assumed to be close to inverse square. The distance d_i is likely to be larger than d_a and the transmitter and receiver are not sited for an unobstructed path. Thus earth curvature and terrain factors will make the inverse square law assumption less inaccurate. However, if such an assumption is made, then the actual receive interfering power level will be less than the assumption so that such an assumption is conservative. Therefore the interference distance can be stated as:-

$$d_i \doteq \sqrt{\frac{G(\theta)\ G(\phi)}{G(0)\ G(0)} \Gamma_1\ d_a^2}$$

or $$d_i = d_a \sqrt{\frac{G(\theta)\ G(\phi)}{G(0)\ G(0)} \Gamma_1}$$

where Γ_1 is a corrected form of Γ to allow for the increased losses on the interference path.

Thus the distance to an interfering transmitter (or to a cochannel receiver from that transmitter) is a simple scaling of the length of the normal link path from that transmitter.

The task of calculating this is still not simple if the true polar diagrams of the aerials are used; so these are approximated to the shape shown in figure 13.3.3 - the keyhole shape. This has only two values unity (or 0dB) in the beam-width (40° in this case), and -8dB elsewhere (this takes account of the worst peak sidelobe from the poorest catalogued performance).

The $\frac{G(\theta)}{G(0)}\ \frac{G(\phi)}{G(0)}$ product therefore has your values depending on the relative directions on the aerials, as shown in table 13.1

TABLE 13.1

$\frac{G(\theta)}{G(0)} \cdot \frac{G(\phi)}{G(0)}$ dB		Is the interfering transmitter in the direction of the main beam of the link receiver	
		YES	NO
Is the interfered with receiver in the direction of the main beam of the transmitter	YES	0	- 8
	NO	- 8	- 16

The last equation for d_i can be written in the form:-

$$d_i = d_a \sqrt{\frac{G(\theta)}{G(0)}} \sqrt{\frac{G(\phi)}{G(0)}} \sqrt{\Gamma_1}$$

The last factor is a constant, the second to last represents the square root of the polar response of the receiving aerial - this has only two values. Taking each in turn of these values then d_i is a function of:- d_a, two constants, and the square root of the polar response of the transmitting aerial. Therefore an outline cut to this shape would, if suitably scaled, represent the distance d_i on a map. Another shape would represent the distance for the other value of receiving aerial gain.

For either of the values of $\frac{G(\phi)}{G(0)}$ the scale is independent of map scale and its linear size is proportional to $d_a\sqrt{\Gamma_1}$ ie since Γ_1 is a constant it is proportional to the length of the designed link path from A to D.

A series of keyhole shapes can therefore be drawn having the form shown in Figure 13.3.4, with the line marked in terms of modified protection ratio Γ_1 determined from:-

$$\Gamma_1 = \left(\frac{d}{d_n}\right)^2$$

If these shapes are in coloured transparent material they can be laid on a map of links (the real version of figure 13.2.2) to determine whether a link receiving site can be assigned cochannel to the link being assessed.

The method is as follows.

Take a transparent keyhole shape of a size such that the interfering link transmitter lies at the centre of the circle and its intended receiver lies on the central line at the desired protection ratio. Lay another transparent shape, of the form shown in figure 13.3.5, and place its apex on the interfered-with receiver and its central line pointing towards that receiver's intended transmitter. This represents the main beam of the receiving aerial polar diagram. If the interfering transmitter is within the second shape, then use the outer edge of the keyhole pattern, if the interfering transmitter is not within the triangular shape then use the inner dotted line of the keyhole (this is drawn 8dB down on the outer shape). If the interfered with receiver is outside the relevant line of the keyhole then the protection ratio will be met, and conversely.

Comparisons with existing reuse patterns showed that a suitable value for the modified protection ratio was 15-18dB, the latter was used as a starting point but relaxed to the former if known factors of topography could be counted on with certainty.

13.3.3 The Refined Keyhole

The simple keyhole shape described in the previous section was found to be unnecessarily restrictive in that the corners of the sector part projected too far. It proved to be a simple matter to plot the true shape for the main beam using manufacturer's polar diagram data. Such a shape is shown in figure 13.3.6 and this was cut-out in coloured transparent material. The shape is now less like that of keyhole and more that of a shoe or footprint.

The term footprint is perhaps best avoided here since it is used in satellite communication to describe the shape of the service area on the earth's surface illuminated by the satellite aerial. In the case under study it is an interference contour and could perhaps have been derived by considering it to be the available field strength, or power, to a receiver from the transmitting aerial. But this would not have tied the size of the pattern to the length of the link path.

The protection ratio scale was determined by $\frac{OA}{OR} = \Gamma^{\frac{1}{2}}$. 3dB steps were marked from 12dB to 21dB.

Having drawn the outer shape of the figure, the inner dotted line, representing the 8dB less gain of the receiving aerial, can be drawn by scaling. This was most easily done using the device shown in figure 13.3.7. Here the outer sloping line A is drawn at any arbitary angle to the vertical and the horizontal distance from the vertical O to A represents the outer of the shape. The inner sloping line S represents the dotted inner line of the shape and is positioned such that along any one horizontal

$$\frac{OA}{OS} = \left(10^{\frac{8}{10}}\right)^{\frac{1}{2}}$$

$$= 2.5 \text{ for the 8 dB ratio}$$

With the reference origin of the shape on the vertical line it was turned until the outer edge cut the line A on the same horizontal as the reference, a mark could then be made above the corresponding S position. By suitably sliding and turning the shape other points could be made on the 8dB inner line and the points joined up.

A similar scaling device was used to produce a range of sizes of shapes corresponding to range of link path lengths.

13.3.4 Use of the Refined Keyhole

The technique envisaged for using the shapes started with the most heavily used link paths ie those bearing 3 or more channels between the same two points. If the shapes showed that another high usage path could be made cochannel then the reuse distance would obviously apply to all the channels on those paths. So another path with the same or similar usage could be sought, tested and, if suitable, assigned to the same frequency group etc. This frequency group - given an alphabetical designation, say J - could contain any number of channels, say n. There was no implication as to the actual channel numbers or frequencies of these channels, merely that they would be the same ones wherever that group occurred. On this basis therefore the assignment strategy made n county-wide assignments at a time.

If some of the paths covered by this technique had less than n channels then that group would not be fully utilised on a national basis. So a visual scan was made of the link map (the real figure 13.2.2) to try to predict a suitable value for n. For example a dense path might have, say, seven channels, but other suitable paths which met the spacing criterion might have only 4, 3 or 2 channels. Thus a choice had to be made, compromising between assigning as many channels as possible at one time by making n large, and on the other hand minimising the occurrances of covering a link path with a group having too many channels thus leading the poor spectral efficiency. For the case cited above, n would probably be chosen as 4, then the 3 channels left to be assigned at the dense site could be catered for by another reuse pattern K, with n=3.

It was envisaged that towards the end of the exercise n would have to drop to unity.

13.3.5 Overlays

Whilst the technique of planning using the keyholes was being developed, it became possible to get useful data-base outputs from the planning computer [the data bases for this computer were constructed in a manner which was considered efficient for its intended mode of operation but this often conflicted with a form which was of direct relevance to the independent planning exercise]. This output gave the frequency (effectively the channel number) of all the return links in use at that time. From these it was possible to produce a histogram of the degree of reuse currently employed. This is shown in figure 12.2.2.

From this figure it will be seen that some channels are reused very many times whilst others have a very poor utilisation. The data-base did not represent the entire situation, it was known that some counties had not been fully recorded, but the implications were significant. If one channel could be reused 13 times then there was every reason to suppose that many others could achieve the same reuse. If an average of just 3 channels were to be assigned, in the manner of the alphabetical group of the last section, to one of the 13 reuses then this would, straight away, effectively assign 39 link channels. Extending this, the cumulative number of channels reused more than 10 times is, from the histogram, 198. If each of these were to cover an average 2.5 channels then 500 channels could be assigned.

The beauty of using existing reuse patterns was, of course, that they had been proven to work over the course of time. Another advantage of this technique was that the speed of allocation could be increased, although the existing reuse patterns did not necessarily conform to a minimum reuse distance with the consequence of less than the highest spectral efficiency. To make use of this method a reuse pattern for one of the high reuse frequencies was drawn on a transparent sheet which overlay the real map of link paths. This produced a pattern of the form shown in figure 13.3.8 which is derived at the same scale, from the actual link paths on figures 13.2.2. Now although only representing one frequency or a channel of current usage, they could each represent a block of channels as for the keyhole technique. So a scan of the number of channels on each link path could give an indication of the number of channels to assign to that block.

13.3.6 Keyholes and Overlays Combined

At this stage of the overlay technique the addition of the refined keyhole criterion gave significant boost to the spectral efficiency.

A visual scan of the link path map (fig 13.2.2), with the reuse pattern overlay on, it could well show that extra paths could be brought into the pattern. Take for instance the link shown as A on figure 13.3.8. This might be only a 2 channel path whereas the majority of other paths could be 4 channel ones. On inspection an associated link path, such as B, might carry only one channel, but if it could be incorporated into the pattern then this extra link path could be covered 'for free'. Otherwise two of the two channels which would be available for path A would be unused.

Whether this path B might infringe the operation of other links could be tested by a suitable size of refined keyhole.

Similarily link path C shown on figure 13.3.8 although not part of the original reuse pattern, could, after test with the refined keyhole as is shown, be included.

13.3.7 Record Keeping

Having assigned the first block of link channels as a group, say group G, it was necessary to ensure that these were removed from the link map and not be involved in subsequent assignments. The technique evolved for this was to keep a master version of the link path map and to reduce the numbers shown against those link paths which had been assigned. Although this sounds simple, even on a large scale map it entailed many crossing outs of numbers and was vulnerable to errors.

So a double check was kept by updating the print-out of link assignments by making alpha numerical assignments to each link. For example if group G covered four channels then on the print-outs, those paths assigned were designated G1, G2, G3, G4. This ensured that they were not duplicated onto another grouping.

13.3.8 The Outcome

Initially 3 and even 4 channel blocks could be assigned, but fairly quickly the majority of high usage paths were assigned. Subsequent reuse patterns then covered only one or two high usage paths and a number of low usage ones - frequently only single channels ones. A judicious use of the refined keyhole was of use here, but ultimately only single channel groupings could be assigned. This, coupled with the record keeping activities, made the task a demanding and labour intensive one. It was therefore time consuming and no method of mechanisation (computerisation) could be devised to shorten the time involved.

Whilst investigating inconsistencies, more and more unrecorded links were discovered and added to the original total. The onward links were catered for in this exercise on the grounds that, may be they could be implemented in practice (given high performance hardware), but if they were not included in the frequency plan then they could not be realised no matter what the equipment performance. The end result was that 640 (not the original 510) linkpaths required allocating and these were accommodated on 72 channels ie 0.9 MHz of spectrum.

Thus the desired objective of housing the links within 56-60 channels was nearly achieved. Due to the time consuming nature of this aspect of the frequency planning task, the others were carried out in parallel with it. The planning tactics changed because of these other activities, but the desire to minimise the spread of the channels associated with the links, or the total bandwidth which they occupied, remained constant.

13.4 Assigning Main Transmit Channels

13.4.1 Objectives

The previous chapter (12) concluded with the aim of assigning all the base station transmit channels in the 154-156 MHz band. To minimise the spread of the associated high order intermodulation products in the receive bands, the occupied bandwidth of the main transmission (and the link transmissions) at any one base station should be minimised. The initial assessment given in that chapter, was based on the assumption of 100 kHz for each, ie up to 8 channels for main transmission and up to 8 for links.

The other factors which must be taken into account are; the chaining phenomenon, and the reuse distance/protection ratio constraint; as given in section 12.1.3(f). The method of dealing with these will form the subject of the succeeding sub-sections.

13.4.2 Chaining - General

Chaining, it will be recalled, is the mechanism by which assignments in one county can be directly affected by those in an adjacent one. This occurs when the counties concerned share a base station, so that, at this location, the assignments pertinent to both counties are now to be ideally contained within 100 kHz (8 channels), or at least in as small a spread of channels as possible.

The next subsection (13.4.3) will discuss the desirability of assuming that all base stations within a county can radiate all that country's channels. The chaining constraint is not however carried over, by this mechanism, to these other sites, ie only the sharing site need carry both counties' channels. Even here, this is a demanding requirement where the sum of the number of channels in the two counties couñcerned is large. If the sum exceeds 7 then the requirement to contain them within 8 channels cannot be met [there needs to be a 'guard channel' unassigned between the county allocations to meet the constraint of 12.1.3(o)]. Whilst this 8 channel band spread constraint is a desire, rather than an absolute limit, it was felt best, at this stage, only to accommodate at the sharing site those channels of each county which are actually required there, rather than the totality for both counties. If it were later found to be desirable to include those other county-wide channels, each one could be assessed individually for the impact of high order intermod products in the receive bands. Since adjacent channel allocations within a county is now assumed to be admissible then the allocations can take the form shown in figure 13.4.1. This shows a channel number assignment, corresponding to an actual frequency, and at the site, number V, the dispositions of the channels for the counties which share the site. These are channels 3 and 1 of county M (their local county numbering scheme, which always starts from one) and channels 1, 2, 3 of county N. Since the local numbering schemes are merely their local designation they are not tied to frequency plan numbers and can be assigned in any order.

The gap of an unassigned channel at channel number 42 is that required to ensure that adjacent counties do not operate on adjacent channels. It might seem strange to keep this requirement when that of not operating on adjacent channels within a county is no longer a requriement. The argument in favour of this is that a mobile which could experience interference from one of its own adjacent channels, because that is of a significantly higher level, can change to that channel. This at least is true in principle, and he might well be advised to do so in practice since the implication is that this adjacent channel is some 50dB stronger than the one he is operating on. If the mobile were to experience strong adjacent channel interference from an adjacent county assignment however, then, he could not change to that channel. Such circumstances are unlikely to be frequent but could happen near the county borders where coverage from an adjacent county's site might be better than that of his own for small regions.

13.4.3 Chaining - Specific

Figure 13.4.2 shows how the other channels for each of the sharing counties might be assigned. The tight grouping for a county's allocation is a requirement since a site may radiate any, or all, if the assignment on a county-wide strategy is followed, of the channels. This figure shows only two counties in a chain; figure 13.4.3 extends this to two more chaining counties.

This figure shows a reason for the odd numbering sequence at site V. Local channel 1 is often the county-wide one and is therefore to be found at both chaining sites in the county. To minimise the spread at any site it is therefore advantageous to locate this in the spectral channel position near to the middle of the county's assignments. Otherwise one or the other of the sharing sites could find it to be an isolated one and thus extending its local spread. Channel 4 of county M also illustrates another feature to be encompassed. This shows that even if a county channel is not involved in either of the sharing sites of that county, then it will still cause band spreading at

one of the sites. It cannot be located elsewhere in the spectrum without causing a band spread problem at some site(s) in the county. The only exception to this would be if it were radiated from one site only and with no other channels; this is extremely rare and in any case breaks the constraint of attempting to allow all the county's channels to be radiated from all their sites.

Branches or loops in the chain also cause bandspreading and a skeleton form of an actual assignment pattern takes the form shown in figure 13.4.4. This is illustrative and does not show the actual local channel numbers, and merely gives the counties and sites each a serial number. Apart from showing the effect of branches and loops the other feature demonstrated is the constraint which chaining still imposes on the assignment pattern. The principal feature of this is that the pattern is in the form of a cascade. It is not possible to reverse the direction of this cascade due to the coupling of counties by their common site. The closeness of counties in the figure reflects to some extent the geographical closeness of these counties. To reverse the direction of the cascade would mean reusing frequencies (channel numbers) closer than is allowed by virtue of the desired protection ratio. Thus the chaining constraint implies that the total number of channels occupied by the counties in that chain cannot be less than the sum of their individual requirements plus the guard channels required between adjacent counties!

The tactics of main transmitter assignment are, therefore, to determine the longest chain (most counties); assign this using the techniques just described; and then to see what can be accomplished by way of frequency reuse for the remaining chains and individual (unchained) counties.

13.4.4 Reuse Distance - General

Techniques for minimising reuse distance were discussed in section 5.6, the frequency reuse aspects of area coverage. This showed the benefits of multiple transmitter operation on an idealised planning basis. By increasing the number of base stations and in consequence lowering their transmitter power the reuse distance could be significantly lowered compared to the single base station case. This technique has to some extent been employed in providing coverage of a county. Several, rather than one dominant, base stations are employed. Thus a certain amount of spectral efficiency has been gained and reuse distance minimise already; to improve on this would mean the deployment of more base stations. This cannot be contemplated, certainly within the time scales of the

WARC frequency change programme; apart from the cost and technical planning load, the planning approval system would take far too long. Thus the reuse distance to be considered here is that from the nearest base station in county 1 to the edge of county 2, where the two counties are operating on the same spectral channel number.

The frequency planning computer discussed in section 12.1.2 contains a protection ratio matrix. This gives a value of protection ratio from any one base station site to any other. This is used as a look-up-table in the relevant part of the constraint sieve to see whether an assignment at a site would interfere with the channel already assigned to a distant county. All base station sites in that county would be included in the computer assessment and, if any one (most probably the nearest) indicated a protection ratio lower than the current level for that constraint, then this would be indicated. This corresponds fairly well with the aim of edge of county assessment given above. The protection ratio figures used in this matrix are derived from CCIR 576-1. This provides a curve relating protection ratio to distance from transmitter for 'typical' types of terrain (with aerial heights and percentage of time as parameters). The difference between using the desirable edge of county receiving position for a mobile and that of using the nearest base station is to some extent covered by taking a judicious allowance for the receiving aerial height. Thus the protection ratio matrix is just a distance matrix in another guise - any value of protection ratio corresponds to a given distance on the typical terrain.

For quick planning purposes therefore the corresponding distance criterion is much easier to use. For the transmit power levels envisaged for the system (see section 10.1), the aerial heights for both transmitters and mobiles, the percentage of time requirement, and using the typical terrain curve, the distance corresponding to the 20dB protection ratio for a 4μVemf (-131 dBW) received signal was 150 km. (This assumes a transmitter aerial height, above surrounding ground; of 300m.)

To assess whether a distant county can operate on the same channel as a reference county, all that needs to be done is to use a pair of dividers, set to the scale equivalent of 150 km, on the national map (figure 13.2.1). In practice, having assigned two or more counties to be cochannel then the dividers need to be used from all the nearest counties. It proved more easy to construct more transparent shapes (as used for link planning purposes) but this time having a semi-circular shape with a radius correspondign to 150 km.

The shape could be laid on the map with its straight edge on the tip of the county already assigned a channel. The tip was used rather than the actual base station to allow for any future redeployment of base stations into the most unfavourable location. If the curved edge of the shape could be made to include the boundary of the county under assessment then this county was not admissable. With two or three shapes, assessment of admissable counties could be made a very rapid process. In fact the insight gained was much higher than would be expected from the equivalent computer aid sieving process, besides being quicker. Very occasionally the 150 km barrier was broken where it was highly desirable to do so for spectrum reuse purposes and where known geographical features would enhance the expected protection ratio. This was only admitted where the 150 km constraint was only just broken and a check was made to ensure that an identical or more stringent reuse pattern existed for the current 100 MHz main transmit bands.

14.3.5 Reuse Distance Specific

Armed with this masking technique the next longest chain would be assigned. Initially a county at one end of the chain was selected and a cascade of channel assignments made as before, but using arbitary channel numbes and drawing the cascade on tracing paper. When completed the tracing paper could be laid on the master channel assignment chart (of the form shown in figure 13.4.4) and juggled for position. The position was determined by the desire to find the closest spacing on the chart which complied with the 150 km reuse distance on the geographical map.

To do this each of the new chain's counties were compared for cochannel assignments by looking along a vertical of the channel number. For the counties which shared channels between the chains, the check was made using the 150 km masks. It would have been trivial had the distance separation between cochannel counties been the same and had the corresponding counties required the same number of channels. In practice not only did this not apply, but the differences in branches and loops were an added hinderance. However, by forcing an occasional extra unused channel into one or other of the cascades a suitable fit could be made.

This then temporarily froze the assignment at that stage and the next longest chain could be considered on that basis. When the chains were complete attention was turned to the non chained counties, thus leaving the easiest and most flexible assignments to last.

Two independent assignment patterns were completed on this basis by two people. One had the property requiring the least amount of spectrum at 88 channels and leaving significant flexibility for manoever, the other was slightly mlore demanding in spectrum at 92 channels and had somewhat les flexibility, but it included the Metropolitan Police forces requirements for 75 channels!

13.4.6 Conclusions

An assignment technique was developed for allotting all the main transmit channels within just over 1 MHz of spectrum. This is well within the available 2 MHz. The extra bandwidth available could be used to provide increased flexibility for future growth or changes. Such flexibility would best be provided at this assignment stage by effectively increasing the number of channels required by each county. The increased band spread at each base station would be small.

The assignment was accomplished within the constraints set.

13.5 Assigning the Link Receive Channels

13.5.1 True Intermodulation Position

The aim, at this stage, was to assign both the link channels and main channels on the base station receive side so that they both lay within the lower receive band of 142-144 MHz. At any one base station the actual assignments must not lie within the high order intermodulation bands. These could now be assessed since the definitive link and main channels for each hill top site were available from the work described in sections 13.3 and 13.4

The section 12.5 showed diagramatically where these intermodulation products would be located for certain specific transmit allocations. It was necessary at this point to be able to explicitly calculate the intermodulation locations for any distribution of link and main transmit channels. The technique for so doing is described in appendix X - High Order Intermodulation Calculations and Results. Chronologically the technique (an adaption of the intermodulation display calculations of Chapter 11) was devised at this point although its results had been anticipated in some of the reasoning of Chapter 12. Explicitly it confirmed the absence of gaps in the receive bands when the transmit channels formed a sequence giving 3rd order intermod free performance for 13 channels. It also showed the number of intermodulation products which fall on each receive channel to be surprisingly high. The effect of the main and link transmit bands, each operating with adjacent channels, corresponded to the predictions shown by figures 12.5.1 - 12.5.3.

Those figures could not, however, predict or show the number of intermodulation products which would fall on each receive channel for this case. It was expected to be a high value when the generating transmitter channels were closely spaced, as in this case, but even so the numbers derived by the calculating technique were staggering. The implications are that, for the 13 transmit channels shown, there are some 300,000 11th order products falling in the receive bands - and this is only one of a number of sets of 11th order products, the others falling in regions of no interest to the current argument.

The appendix discusses the implications of so many intermodulation products on one channel and concludes that they must be deemed to add on at least a power basis. Thus, where there are say 300 of them, then this will increase the interference level by 25 dB compared to the case of just two transmitters. This adds weight to the arguments for avoiding these channels for receiver allocations.

13.5.2 The Number of Channels Required

Section 13.3 described the extensive work required to ensure that the transmit links could be allocatted less than 1 MHz of total bandwidth. It required a combination of existing reuse patterns and enhancement using a keyhole shape. The task for the receive side is, however, extremely easy. Again the requirement was to occupy no more than 1 MHz of spectrum in total (leaving 1 MHz for the main receive channels). Present usage was such that only 80 channels were needed to implement the national needs. Therefore the current assignment could be taken, (in the 146-148 MHz band) and the lowest of those channels assigned to the lowest in the new band; the second lowest in the present bands to the second channel of the new band; and so on. Thus packing into 1 MHz of spectrum as desired.

This did not, however, entirely meet the requirement for assignment, since no account of avoiding the intermod channels had been taken. In fact, there was no real guarantee that if a complete 1 MHz of the lower receive band were allocated to the receive links, then at any one site the portion of the spectrum allowable from cochannel re-use considerations would not have clashed with that blighted by the intermodulation products. This chance might have been small, and overcome, for the occasional specific case, but it did sound a warning note - one which was to be taken up and amplified by the considerations of the next section.

13.6 Assigning the Main Receive Channels

13.6.1 Numbers of Channels

As was found with the link receive situation, there were fewer channels in use at that time for the main receive, compared to the main transmit. The number was 76, so that they also could fit within the desired 1 MHz of total bandwidth.

It is worth commenting on the difference in the number of main receive channels compared to the main transmit. It implies greater re-use of the receive channels which appears to be due to a combination of; absence of chaining constraint, and slightly shorter tolerable re-use distances. Why the latter should be admissable is not clear since arguments can be advanced for the go and return directions to be balanced in this respect. However, the situation was advantageous and was defendable on the grounds of current usage.

13.6.2 Actual Assignments - a Fly in the Ointment

The stage was now set for the mere task of making the detailed channel assignments for both the main and link receive positins. [Incidentally, it might be of interest to disclose that a date had been set, less than a week away, for this frequency plan to be presented to interested parties within the Directorate]. Problems, which should perhaps have been foreseen, but which were not, were then disclosed.

The view presented so far for the receive assignments is virtually just that of avoiding the (narrow) bands of intermod products at any site. But in fact, it was necessary to consider a whole county. For this county, all the base stations must have the same assignments for the main receive channels and for the link receive channels. (The links will all be the same since it will be recalled that there is only one outgoing link assignment from control for each mobile channel, and this forms the base station receive channel). Similarly, all the base stations will radiate on the same frequency assignments throughout the county. But the link transmission from the base stations will all be different!

Thus, although from figures 12.5.1 - 12.5.3 it would appear that the bulk of the lower 2 MHz receive band was available for receive assignments, in reality the intermods from all the sites, which will appear in different regions of the 2 MHz, must all be considered together. This is equivalent of overlaying the

intermod swathes of figures 12.5.1 - 12.5.3 on top of each other (and others to the number of base stations in the county) and then seeking room for receive assignments.

Had this been realised during the work described in Chapter 12, then a different strategy would have been adopted since it violates the conditions, set there, of requiring that there was a very high probability of the detailed assignments being realisable. For a four station county there will be four swathes of intermod bands across the chosen receive band. Even if the county operates with only a few channels the total effect looks most unpromising for the assignment of suitable receive channels. It would be at best a tedious task.

There was, of course, a lesson to be learnt here - that is that the details of a complex task are of great importance and can nullify the optimism given by a broad overview of the situation. It was in fact a lesson thought to have been learnt and the only excuse was that the bulk of the work described in Chapters 12 and 13 would have to have been done to disclose it.

13.6.3 Less Spectrum Available - No Ointment!

Before this factor had finally been assimilated and thoughts developed for coping with the situation, another, and at first sight bigger, problem presented itself.

Throughout Chapter 12, and so far in Chapter 13, it had been assumed that the band 142-144 MHz was available for our assignments. This was initially a working assumption, based on certain beliefs, and one which quickly was taken as firm. But these initial beliefs were discovered to be not only unreliable but in fact erroneous - only the band 143-144 MHz was available for receive (together with the band 146-149 MHz discarded in Chapter 12).

Thus a large portion of the assignment strategy had been based on incorrect assumptions as to the bands available - a simple numerical error!

13.7 Reassessment of Strategy

13.7.1 A Review

The frequency planning saga is about to take another twist, bringing the argument round into another loop of the coiled path. It is, therefore, sensible to take stock of the position at this stage. This is shown in figure 13.7.1

This shown the dispositions of the block assignment for the four blocks of main receive, link receive, link transmit, main transmit. The bands (erroneously) chosen for the receivers would have some intermodulation products of 11th and 9th order. Furthermore, the assumption that these intermodulation clusters would be small, and therefore leave room for actual receiver assingments free of them, did not stand up to more detailed analysis on a county-wide basis.

One or both of the receive allotments will obviously have to be relocated in the 146-149 MHz portion, but the intermodulation products will be of lower order here, hence stronger, and not tolerable in general.

13.7.2 A Possible Solution

The statement that one or more of the receive allotments would have to change could be refuted. But the arguments which lead to the disposition of the transmit bands were no longer entirely valid. The use of the 152-153 MHz band was brought about by the desire to make 3rd order free transmit assignments - even if only for the main transmitters. But this was shown in section 12.5 to be incompatible with avoiding high order products in the receive bands, and the latter factor was considered to be dominant. Since 3rd order free assignments were no longer the aim, then the usage of the transmit bands could be reconsidered.

A simple solution then presented itself. The total bandwidth required by the main and link transmit channels was 92+68=160 channels; just that which would fit into the 2 MHz band from 154 to 156 MHz. On this basis there would be, at any one base station, a bunch of channels assigned from the main transmit band and a group of link channels from the link transmit band - hopefully a narrow group. One of the receive allotments could be in the band 143-144 MHz and the other in 146-147 MHz. The situation would be as depicted in figure 13.7.2.

The regions blighted by high order intermodulation products were then those shown in figure 12.1.1. This shows that the upper of the receive bands would suffer from 11th and 9ths order products. Due to the bunching of main transmitter assignments, and the hope for bunching of the link transmitter ones, there was some hope for finding gaps in the intermod swathes at any one base station. Even if this were not true then by putting the receive link band in the position shown, only that would suffer. Since the link receive level is designed to be significantly higher than the main receive level, then the impact of intermods would be least for this situation and the situation could then be considered, if not ideal, tolerable.

If indeed it were deemed acceptable to operate the link receivers in the presence of 9th and 11th order products, then both the main and link transmit allocations could be freed from the bunching constraint. A more detialed examination of this hypothesis showed, however, that this was true for the links but not true for the main transmissions. The chaining phenomenon coupled with the need to avoid adjacent channel allocations in adjacent counties and the overall constraint on total bandwidth occupied prevented this.

Thus some recovery was made to what had appeared to be a very difficult position.

13.7.3 A Better Solution. I

Further thought was given to the situation to see if the solution just described could be improved. A quick assessment of the transmit linking situation showed that, even if a bunched channel assignment (ie adjacent channels or very close to that) were made at one site, then it would be very difficult to ensure that more than some 3 or 4 could be so treated. The others were dependant on the allocation of the first and subsequent ones and were also interrelated. Therefore, for most sites the link band allocatios must be expected to fall anywhere within the link transmit band. This immediately implies an extensive spread of the intermodulation products.

Since the receive channels are, however, clustered, then with the arrangement of the bands shown in figure 13.7.2, only those main transmit allocations which fall in the lower part of its band (154 to approx 155 MHz) will in fact cause the high order products to lie in the link receive band (146-147 MHz). The lowest main receive frequency that will just clear this band of 11th order products can be calculated.

Figure 13.7.3 shows the situation. For a band of width x MHz, assumed to have a continuous distribution of transmitters, then the Nth order products will extend to $\frac{1}{2}(N-1)x$ MHz below the bottom edge of the band, or $\frac{1}{2}(N+1)x$ MHz below the top edge of the band. So for 11th order 6x = 9 MHz, ie x = 1.5 MHz, so that any site with assignments entirely above 154.5 MHz will be entirely clear of damaging intermodulation products on the receive band.

No way could be seen of ensuring that all the assignments country-wide could be squeezed into this total bandwidth for main and link transmitters. But it did show that only about half the sites were in fact likely to suffer from the high order intermod effect.

This line of thinking then led to consideration of placing the link transmit band somewhere in the middle of the 154-156 MHz band. If the main transmit band could be split into two portions so that at any one site assignments were made in either one or the other (together with any assignments from the link band), then again it should be possible to ensure that the intended receive bands were completely free of intermods to the required degree. The situation is illustrated in figure 13.7.4a and b.

The aim is to determine r and s such that any transmissions in the band x or y cause no 11th order products below 147 MHz. From the reasoning given above x = 1.5 MHz and therefore r = 0.5 MHz. Similarly, the above formulae can be used for the second case such that:

$$5y = 154-147 = 7 \text{ MHz}$$

Therefore y = 1.4 MHz and s = 0.6 MHz.

This means that a 'core' band t of width 2-r-s = 0.9 MHz (72 channels) can be located with its bottom edge at 154.5 MHz and can be used freely anywhere in the county with either; assignments solely in the bottom portion left of the 154-156 band; or assignments only in the top portion of the band. This is shown in figure 13.7.4c. This core band is not synonomous with, or identical to, the link transmit band. Since the link band requirements are narrower than this, then the link band can be located anywhere within the region t. The unoccupied portion of t can be made part of the relevant upper or lower main transmit band.

Thus, if the chaining constraint, which required a contiguous band 92 channels wide from section 13.4, could be broken by breaking the longest chain near its middle, then the desired objective could be achieved. To do this would entail either providing two (not too close) base stations in place of the one common one where it was desired to break the chain, or simply to declare that one of the two counties involved at that site could not use the channels radiated from that site anywhere else in the county. For the sake of avoiding high order intermodulation product interference throughout the nation then, this would be acceptable. It also meant that the main transmit assignment plan would have to be split at a frequency which ideally did not split the assignments at any other site, or for any other county. This proved easy to satisfy.

It would give very great flexibility, since no intermodulation considerations need be made for the detailed assignments of any of the three full bands and two sub bands as shown in figure 13.7.5. The only constraints would be adjacent county operations and re-use factors. Thus the requirement for growth, and more particularly flexibility for changes, are well met.

13.7.4 A Better Solution II

The one real disadvantage of the plan proposed in the previous section (apart from the 3rd order intermod aspects, of course), was that of having to break the chain of counties sharing common base stations. It was chronologically then some two days before the presentation of the plan was to be made (see 13.6.2) and the salvage work which had been necessary had in fact produced a better plan that that originally envisaged. This was partly due to the work which had been in train to minimise the spectral spread of the transmit bands. Since this had been an outcome of the erroneous assumption of extra receive bandwidth, then that assumption could be viewed as beneficial.

At this stage the observation was made (Pearce 1983) that in fact the chain need not be broken. All that was required was to cut the previous main transmit plan (section 13.4) into two approximately equal halves, and put them on either side of the link band. Due to the bunching of main transmit channels one of the bunches would be cut in two. Thus at one site the main transmissions would straddle the link transmit band. If no other allocations were made at that site, other than in the bunch or link bands, then again the 11th order intermodulation products would not reach the upper portion of the receive bands. The situation is shown in figure 13.7.5. The link band is drawn as 72 channels wide (0.9 MHz) and the total number of channels covered by the main transmission at that site is assumed to be 16 for the sake of illustration (a figure which is most unlikely to be reached). Thus the total possible spread of transmissions from that site is 88 (1.1 MHz) and it will be seen that the nominated link receive band is well clear of intermodulation products.

Thus this version of the plan (figure 13.7.6) accommodates all the requirements it was found possible to meet. It has the very large measure of flexibility described in a Better Solution I (section 13.7.3) in that assignments can be made anywhere within the bands or sub bands shown. The only exception being at the site where the assignments straddle the link transmit band, and even here there is a significant area for movement before the high order intermodulation factors could start to affect the link receive band.

13.8 Conclusions

The form of the frequency plans devised is given in figure 13.7.6, with only the main transmit band(s) requiring any real comment on their structure. This structure is that of figure 13.4.5 but cut vertically near the middle of the diagram in such a way that the assignment at one site only, is split, and appears in both main transmit sub bands. The detailed assignment of numbered channels to actual base stations is merely a (tedious) exercise, given this structure and the rules developed in this chapter for such assignments. The details are neither necessary to this document nor available under security considerations.

The topic of security allows one other feature of the plan to be discussed in passing. There are a number of particular assignments within these frequency bands which are for various special services and which do not necessarily conform to the county structure. Due to the flexibility of the plan devised they can easily be accommodated.

The objective set at the start of Chapter 12 was to see if any viable plan could be produced. The answer has been shown to be yes - provided one major and one minor constraint can be broken. These are transmit channels to operate in a non 3rd order intermodulation free environment, and onward links are not catered for. These, it was said, must be relocated elsewhere (microwave bands?). Having said that these constraints must be broken, there may be some room, due to the plans flexibility, to restore at least a measure of these.

3rd order free assignments could probably be made for at least a few main transmissions. These may not give any benefit initially due to the non 3rd order free assignments within the link transmit bands. But, if and when, those links are relocated (at microwaves?) then the main transmission sequences will work. When this happens, however, the whole of the link band will be unoccupied and it may well be best to consider total main transmit re-assignment at that stage! Likewise it was suggested, at the presentation, that it might be possible to accommodate the onward links in the unused 147-149 and 152-153 MHz bands. Again the answer must be one of, heavily qualified, possibly. Such usage will depend on realisable filters and the intermodulation products produced by the use of such bands.

In summary then, a frequency plan was produced in a very tight time scale and the techniques underlying this plan have been described here. It is best regarded as a 'base line plan' or a 'fall back plan', that is one which is probably capable of improvement or even of being overtaken by the work of others, nevertheless the objectives were achieved.

The route taken from the start of the exercise (the start of chapter 12) has been a tortuous one with frequent backtracking to a previous decision point and a new decision taken. With hind-sight (or perhaps experience) the final form could, or perhaps should have been, obvious earlier. But had this been true, and the final form presented as such, then others would have raised, in time, all the questions of the form "but what if we tried...". At least by following the route described then most, if not all, of such questions have been taken into account.

It must also be pointed out that this exercise was completaly self-contained within the Directorate's 140/150 MHz bands. This required consideration of the odd order intermodulation products only. No account was taken of other users who may share the site, and in particular, the tower. Nor was the impact of working on the 70/80 MHz bands, which will be present on most of these sites, covered. Such factors were not included in the task set but will have to be considered, in the complete design of the system, thus this reported work can be taken as one of the early traverses of the interaction diagram of section 3.2

13.9 Implications

A number of implications, ranging from technical to philosophical, emerge from this exercise and are worth recording - in no particular order of importance:

-Attention will have to be paid to the specification for the mobile equipment to tolerate/overcome having to operate in a non 3rd order intermod free environment.

-Research is needed into techniques for overcoming or tolerating the interference received in this regime.

-Fixed link paths must be installed, and transmitter power adjusted, to give just the design value of received signal level.

-Only 4 MHz out of a total of over 7 MHz has been used to produce this plan. The unused sections are considered virtually unusable - what reasoning went into the acquisition of the spectrum?

-This plan was produced in less than 2 months of intensive work using virtually only maps, shapes, and large tables to work on. Were consultants and a mini computer necessary?

-Since the intermodulation products of the mobile receiver and those caused by the site nonlinearities are either avoided or tolerated, then there are implications for base station equipment design. This could perhaps have relaxed specifications - though the 70/80 MHz and other users will have to be borne in mind.

-All the fixed links will have to have their frequencies changed - twice! This will have to be completed before the main WARC frequency change!

-The onward repreat links that cannot be accommodated must be reallocated before the start of the main frequency change.

14. Conclusions

14.1 General

This thesis has discussed the design of a VHF mobile radio system for the United Kingdom emergency services which had to be completed against numerous constraints. The philosophy adopted was, as far as possible, to set the scene for each major design aspect by way of considerations applicable to mobile radio in general, and to follow this with theoretical discussions which quickly lead to the need for objectives to be restricted to this specific case.

A feature of the thesis was to show that the theoretical aspects were of value only upto a point where further theoretical work would be unprodcutive for the particular task considered. A dominant factor here was that of having to produce a workable design in a very short time scale, and which would only be fully provable at the time of later installation. Thus theoretical aspects had to give way to consideration of:- practical trials, evaluation and measurements, time factors, costs, availability/realisability, and the availability of implementation resources. The design process was shown as a flow chart which was termed an interaction diagram since it related the design elements which interfaced with each other, and it had the form of an inverted tree with the roots being the defined starting factors of; grade of service, frequencies of operation, and the area to be covered. The trunk covered the core features such as numbers and location of base stations, type of area coverage system and the modulation to be used. There were three main branches involving the design of:- the base station, the mobile equipment and the fixed links. The interconnections between interacting elements formed numerous feed-back loops and the concept was introduced of design completion being reached when all paths and loops could be traversed without meeting a discontinuity.

14.2 Specific

Although the thesis was intended to cover all aspects of the design, and to provide a record of the reasons for the decisions which were taken, some aspects are worthy of highlighting for their novel contributions.

The chapter on area coverage showed that several base stations were needed to provide radio cover for the typical operational area of a geographic county. The form of operation which made least demand on the frequency spectrum was shown to be common channel, and, in particular, that the transmitters should operate in a quasi synchronous mode. The multiple transmitter common channel operation was extended through theoretical considerations to demonstrate that this technique has the potential for great spectrum economy, which could be of particular interest to the currently development cellular radio systems.

The question of choice of modulation was examined in depth. Of the three main contenders; AM, FM and SSB, the least used and least fashionable, AM, was shown to be the best choice for the constraints considered. This area was a particular one where practical trials and demonstrations were necessary not only to back the theoretical prediction, but also in order to provide design parameter limits.

The subject of base station design showed that one topic was to become of major importance. That was the influence of intermodulation products generated by the possible non linear aspects of the sites themselves. The so called rusty bolt effect was shown to be present at levels which mean that 11th order products could be of a level to significantly affect the base station receiver. It was expected that this would have a severe effect on the production of a frequency assignment plan for this particular system, but it will also be of great concern to other operators particularly as the number of channels radiated from a site increases. This increase is likely from two causes, the general growth in mobile radio operations, and the desires of planning authorities to minimise the number of base stations by making intensive use of existing ones.

The frequency bands chosen for operation, for in fact they became part of the design process since none were readily available and therefore they had to be sought, were less than ideal in extent and relative positions. The chapters on frequency assignment showed the extreme difficulty of this task for the many layers of constraints which existed. There was severe interaction between the fixed links and the mobile assignments since in the new situation they shared the same narrow allocations. It was shown that it was necessary to break one of the traditional constraints in order to force a solution. Even so, all the existing fixed links had to in effect be replanned - a major task in itself. Novel techniques for doing this quickly were developed which were based on the use of transparent map overlays showing the frequency re-use possibilities. Using this technique re-use factors of 14 were readily available for fixed links in England and Wales. An allied technique was used for the mobile frequency re-use patterns but here the re-use factor had a normal maximum of four.

14.3 Further work

The intended design exercise was effectively completed. Not all the details were defined (and in this thesis only the more major designs parameters are enumerated), but the outstanding factors are of a minor nature. These can be determined within the design element without repercussions to interfacing elements.

The design task is however never complete since improvements are always possible and more importantly users requirements' change to say nothing of developing technology. It is hoped that this work will at least provide a record of the reasoning which went into the present design so that it can be used to test any future changes for their compatibility or impact.

There are several major areas where further work could be profitable and which would have been followed, time and resources permitting:-

- The implementation of multiple transmitter small area coverage schemes for spectrum economy.

- Development of SSB modulation for use by mobiles - although its use may be limited to the VHF band rather than the developing UHF bands.

- The understanding of passive intermodulation generating mechanisms, the ability to predict levels (it appears that this will have to be on a statistical basis), and above all a cure for them.

- Optimisation of mobile receiver performance when working in a quasi synchronous mode

- Optimisation of mobile receiver performance from the 3rd order intermodulation viewpoint.